



**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS DA EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
INFORMAÇÃO**

LEILA CRISTINA WEISS

**RELAÇÕES SEMÂNTICAS EM TESAUROS:
Um Estudo da Abordagem Pragmática**

DISSERTAÇÃO DE MESTRADO

**Florianópolis
2014**

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
INFORMAÇÃO

Leila Cristina Weiss

**RELAÇÕES SEMÂNTICAS EM TESAUROS:
Um Estudo da Abordagem Pragmática**

Dissertação apresentada à Banca Examinadora do Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências da Educação da Universidade Federal de Santa Catarina, como requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

Orientadora: Prof. Dra. Marisa Bräscher

Área de Concentração: Gestão da Informação

Linha de Pesquisa: Organização, Representação e Mediação da Informação e do Conhecimento

Florianópolis
2014

**Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.**

Weiss, Leila Cristina

Relações semânticas em tesouros : um estudo da abordagem
pragmática / Leila Cristina Weiss ; orientadora, Marisa
Bräscher - Florianópolis, SC, 2014.

166 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina. Programa de Pós-Graduação em Ciência da
Informação.

Inclui referências

1. Ciência da Informação. 2. Relações semânticas. 3.
Semântica. 4. Pragmatismo. 5. Tesouros. I. Brascher, Marisa.
II. Universidade Federal de Santa Catarina. Programa de
Pós-Graduação em Ciência da Informação. III. Título.

LEILA CRISTINA WEISS


**RELAÇÕES SEMÂNTICAS EM TESAUROS: UM ESTUDO DA
ABORDAGEM PRAGMÁTICA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em
Ciência da Informação do Centro de Ciências da Educação da
Universidade Federal de Santa Catarina em cumprimento a requisito
parcial para a obtenção do título de Mestre em Ciência da Informação.

**APROVADA PELA COMISSÃO EXAMINADORA
EM FLORIANÓPOLIS 14 DE OUTUBRO DE 2014**

Prof. Vinicius Medina Kern, Dr.
Coordenador do Curso


Profª. Dra. Marisa Brascher – PGCIN/UFSC (Orientadora)


Prof. Dr. João Batista Ernesto de Moraes – PGCIN/ UNESP


Profª. Dra. Ligia Maria Arruda Café – PGCIN/UFSC


Prof. Dr. Adilson Luiz Pinto – PGCIN/UFSC

AGRADECIMENTOS

À minha orientadora, Prof^a. Marisa Bräscher, pelo apoio, confiança, contribuições e discussões enriquecedoras que permitiram a construção deste trabalho.

Ao Dorlan, pela compreensão e apoio.

À minha família, pelo apoio em todos os momentos.

À Prof^a. Lígia Café, pela confiança e apoio.

À Sabrina, secretária do PGCIN, pela atenção e presteza em todas as situações.

Aos membros da banca de defesa da dissertação, pela leitura atenta e sugestões.

Aos professores do Departamento de Ciência da Informação da UFSC.

Acredito que se há algum Deus, ele não estaria nem em você nem em mim, mas nesse espaço que existe entre nós. Se há algum tipo de magia no mundo, ela deve estar na tentativa de entender e compartilhar algo com alguém. Sei que é praticamente impossível conseguir. Mas e daí? A resposta deve estar na tentativa.

(Celine, no filme *Before Sunrise*)

RESUMO

Analisa como a abordagem semântica pragmática é ou pode ser adotada para o estabelecimento de relações semânticas em tesouros. Utiliza a técnica de análise de conteúdo para caracterizar a abordagem pragmática e identificar o grau de compatibilidade dessa com as recomendações das normas para elaboração de tesouros ANSI/NISO Z39.19 (2005), ISO25964-1 (2011) e ISO25964-2 (2013). Identifica ideias compartilhadas pelos diferentes autores dos textos que compõem o corpus da pesquisa, formado por 79 trabalhos. Agrupa as ideias comuns em 9 enunciados que destacam aspectos importantes do paradigma pragmático no campo de organização do conhecimento e identifica a compatibilidade dos aspectos que caracterizam essa abordagem com as recomendações das normas para elaboração de tesouros. Identificou-se que apesar de alguns aspectos das normas serem compatíveis com os princípios da abordagem pragmática, esses possuem menor importância, uma vez que não se configuram em elementos centrais para o desenvolvimento de um tesouro. Por outro lado, outros aspectos incompatíveis se configuram em elementos centrais para o desenvolvimento de um tesouro, por exemplo a definição de conceito como unidade de pensamento e uma maior preocupação com a representação (essencialismo) do que com a efetiva comunicação. Uma vez que as normas apresentam as práticas já consolidadas no desenvolvimento de tesouros, podemos dizer que a abordagem pragmática ainda é pouco utilizada para o desenvolvimento de tesouros e para o estabelecimento de relações semânticas. Assim, procuramos mostrar como a abordagem pragmática na OC pode ser adotada. Para isso sistematizamos um conjunto de orientações para o estabelecimento de relações semânticas em tesouros, com base nas recomendações das normas e nos enunciados da abordagem pragmática na OC. A variedade de aspectos que caracterizam essa abordagem trazem contribuições distintas para os diferentes tipos de relações semânticas em tesouros. Para as relações associativas destaca-se a importância da bibliografia para a identificação das relações que podem guiar o usuário a localizar informações úteis ao alcance de seus objetivos. O que não se aplica de forma tão efetiva no caso das relações de equivalência, nas quais a identificação da variedade de termos que podem representar um conceito é essencial, e para isso os usuários, e não apenas a bibliografia, são fontes importantes para a coleta das variações linguísticas. Os aspectos contextuais envolvidos na construção do conhecimento, que

são múltiplos e heterogêneos, acarretam implicações éticas para o estabelecimento de relações semânticas em tesouros. A importância de considerar a prática para a formação dos significados pode ser útil para o desenvolvimento de tesouros mais orientados para a comunicação do que para a representação do conhecimento. O entendimento de que os significados variam de acordo com o contexto também é importante para as relações hierárquicas, uma vez que se forem estabelecidas apenas a partir de premissas universalistas podem se tornar demasiadamente rígidas. Ao adotar princípios da abordagem pragmática, os tesouros podem contribuir para uma comunicação efetiva, ao possibilitar a compreensão dos diferentes pontos de vista presentes numa comunidade discursiva e não optar por um como sendo o melhor para representar uma realidade que é, por natureza, heterogênea. As relações semânticas poderiam desempenhar papel crucial nessa comunicação, ao permitir que o usuário possa partir do seu ponto de vista inicial e ir até aqueles defendidos pela frente de pesquisa sobre o tema.

Palavras-chave: Relações semânticas. Semântica. Pragmatismo. Tesouros. Organização do Conhecimento.

ABSTRACT

It examines how the pragmatic semantic approach is or can be adopted for the establishment of semantic relationships in thesaurus. It uses the content analysis technique to characterize the pragmatic approach and identify the degree of compatibility of this with the recommendations of standards for developing thesaurus ANSI/NISO Z39.19 (2005), ISO25964-1(2011) e ISO25964-2(2013). It identifies shared ideas by different authors of the texts that make up the research corpus, consisting of 79 works. It groups together common ideas in 9 statements that highlight important aspects of the pragmatic paradigm in the knowledge organization area and it identifies the aspect compatibility that characterizes this approach with the recommendations of standards for developing thesaurus. It was found that although some aspects of the rules are compatible with the principles of pragmatic approach, they are less important, since they do not constitute in core elements for the development of a thesaurus. On the other hand, other incompatible aspects constitute in core elements for the development of a thesaurus, for example the definition of concept as thought unit and a greater concern for the representation (essentialism) than the actual communication. Once the standards have already established practices in the development of thesaurus, we can say that the pragmatic approach is not widely used for developing thesaurus and to establish semantic relationships. So we try to show how the pragmatic approach can be adopted in the OC. For that, we systematized a set of guidelines for establishing semantic relationships in thesauri, based on the rule recommendations and in the pragmatic approach listed in the OC. The variety of aspects that characterize this approach brings distinct contributions to different types of semantic relationships in thesaurus. For associative relationships, one highlights the importance of literature to identify relationships that can guide the user to find useful information to reach his goals. This is not applied so effectively in the case of equivalence relationships, in which the identification of the variety of terms that can represent a concept is essential, and for this the users and not only the bibliography are important sources for gathering linguistic variations. Contextual aspects involved in knowledge construction, which are multiple and heterogeneous, cause ethical implications for the establishment of semantic relationships in thesaurus. The importance of considering the practice for the formation of meanings can be useful for the development of thesaurus more targeted

towards communication than towards knowledge representation. The understanding that meanings vary according to the context is also important to hierarchical relationships, because if they are established only from universal assumptions they may become too rigid. By adopting pragmatic approach principles, thesaurus may contribute to effective communication, by enabling comprehension of different points of view present in a discursive community and not choosing one as the best to represent a reality that is inherently heterogeneous. Semantic relationships could play a crucial role in this communication by allowing the user to go from his initial point of view to those advocated by the research front on the topic.

Keywords: Semantic relationships. Semantics. Pragmatism. Thesaurus. Knowledge Organization.

LISTA DE ILUSTRAÇÕES

Figura 1 - Linha do tempo das normas para elaboração de tesauros.....	41
Figura 2 - Exemplo de relação USE +.....	133
Figura 3 - Teste todos-alguns.....	136
Figura 4 - Teste alguns-alguns.....	137
Figura 5 - E exemplo de poli-hierarquia.....	139
Figura 6 - Exemplo de node labels.....	139
Figura 7 - Relação derivacional.....	142
Figura 8 - Relação derivacional.....	142

LISTA DE QUADROS

Quadro 1 - Levantamento bibliográfico.....	45
Quadro 2 - Síntese de como a abordagem pragmática se encontra nas normas.....	128
Quadro 3 - Tipos de relações semânticas em tesauros.....	132

SUMÁRIO

1 INTRODUÇÃO	19
1.1 OBJETIVOS	21
1.1.1 Objetivo geral	21
1.1.2 Objetivos específicos	21
1.2 JUSTIFICATIVA	21
2 REFERENCIAL TEÓRICO	23
2.1 PRAGMATISMO, PRAGMÁTICA OU PARADIGMA PRAGMÁTICO EM SEMÂNTICA?	23
2.2 SEMÂNTICA NA ORGANIZAÇÃO DO CONHECIMENTO E NA RECUPERAÇÃO DA INFORMAÇÃO	27
2.3 TESAUROS	32
2.4 RELAÇÕES SEMÂNTICAS	36
2.5 NORMAS PARA ELABORAÇÃO DE TESAUROS	40
3 PROCEDIMENTOS METODOLÓGICOS	45
4 RESULTADOS	49
4.1 ABORDAGEM PRAGMÁTICA	49
4.1.1 A informação é um fenômeno construído	49
4.1.2 Contextos múltiplos e heterogêneos e dilemas éticos	52
4.1.3 A qualidade da informação não é constante	64
4.1.4 A prática é importante para a construção do significado	68
4.1.5 A interação é importante para a validação	75
4.1.6 A bibliografia como fonte para identificar pontos de vista distintos	77
4.1.7 Comunicar é mais importante do que representar	82
4.1.8 Guiar para uma escolha informada	88
4.1.9 Relações semânticas para contextualizar os pontos de vista distintos	94
4.1.10 Síntese conclusiva	102
4.2 A BORDAGEM PRAGMÁTICA NAS NORMAS PARA ELABORAÇÃO DE TESAUROS	103
4.2.1 ANSI/NISO Z39.19:2005	104
4.2.2 ISO25964 – 1	118
4.2.3 ISO25964 – 2	124
4.2.4 Síntese da abordagem pragmática nas normas	127
4.3 ORIENTAÇÕES PARA O ESTABELECIMENTO DE RELAÇÕES SEMÂNTICAS	132
4.3.1 Relações de equivalência	132
4.3.2 Relações hierárquicas	135

4.3.3 Relações associativas.....	141
4.3.4 Mapeamento das relações entre diferentes vocabulários.....	144
5 CONSIDERAÇÕES FINAIS.....	149
REFERÊNCIAS.....	153
APÊNDICE A - Bibliografia Consultada.....	157
ANEXO A – TABLE 4 – Comparison of Mapping Implementation in the Conversion of Index Terms and Search Terms.....	165

1 INTRODUÇÃO

Um importante componente dos tesouros são as relações semânticas. Khoo e Na (2006) apontam a existência de um interesse crescente no estudo de tais relações. Para Hjørland (2003, p. 87, tradução nossa¹), “a unidade básica na organização do conhecimento é a relação semântica entre dois conceitos, e estas relações estão incorporadas/embutidas em teorias”. Essas teorias podem representar pontos de vistas distintos e, assim, apresentar diferentes relações semânticas entre os mesmos conceitos. As dificuldades que se apresentam para representar essas diferenças levam a certo consenso entre pesquisadores da área em considerar o estabelecimento de relações semânticas em tesouros um processo arbitrário.

A arbitrariedade da seleção dos agrupamentos e relacionamentos não pode ser vista, no entanto, como algo totalmente aleatório. Ela se deve aos aspectos que desejamos destacar numa determinada representação do conhecimento. (CAFÉ; BRASCHER, 2011, p. 26).

Deve, portanto, haver correspondência entre o sistema e a realidade que este representa.

Essas e outras questões são estudadas na área de OC, que é interdisciplinar e possui ligação com a área de semântica (estudo do significado). De acordo com Hjørland (2007b) as abordagens diferentes para OC implicam diferentes pontos de vista sobre semântica e ressalta que Peregrin (2004) destaca dois paradigmas dominantes em semântica, o pragmático e o positivista. No paradigma positivista a linguagem é retratada como um meio de representar o mundo e a realidade. No paradigma pragmático, a linguagem é vista como um meio de interação e o significado de uma expressão não é a entidade representada por ela, mas sim a sua função na interação. A abordagem positivista baseia-se no pressuposto de que a linguagem é essencialmente um sistema de nomes para representações de coisas, enquanto que na abordagem pragmática assume-se que a linguagem é uma coleção de meios para fazer coisas. (PEREGRIN, 2004).

¹ The basic unit in KO is the semantic relation between two concepts, and such relations are embedded in theories (HJØRLAND, 2003, p. 87).

Observa-se que a área de OC sofre também influência dos paradigmas positivista e pragmático. Vega-Almeida, Fernández-Molina e Linares (2014) apontam que os pressupostos teóricos pragmáticos também estão presentes no paradigma social em CI. Tanto a abordagem pragmática quanto a social em CI enfocam aspectos sócio-culturais e contextuais envolvidos nos processos de comunicação do conhecimento.

Na OC o tipo de abordagem influencia as características e funções de um SOC. Diante de uma perspectiva em que a linguagem é vista como um meio de interação, e o significado, como a função que uma expressão exerce na interação, os SOC devem também ser instrumentos de interação que levem em conta os aspectos contextuais em que se inserem.

Os tesouros são tipos de SOC e possuem normas internacionais que orientam sua elaboração e que apresentam algumas práticas já consolidadas na OC. As recomendações das normas, de modo geral, são baseadas em processos utilizados para o desenvolvimento de tesouros já existentes e considerados eficientes. A eficiência ou ineficiência de um tesouro depende dos propósitos que são atribuídos a ele. Assim, as recomendações das normas, ainda que não indiquem de forma explícita seus pressupostos teóricos, podem indicar em qual abordagem de OC se inserem.

Diante da influência que a área de semântica exerce na OC, da importância das relações semânticas para a estrutura conceitual dos tesouros, bem como demais Sistemas de Organização do Conhecimento (SOC), nesta pesquisa procuramos responder a seguinte questão: a abordagem semântica pragmática é ou pode ser adotada para o estabelecimento de relações semânticas em tesouros?

Visando responder essa questão definimos os objetivos dessa pesquisa, que são apresentados na próxima seção desse capítulo de introdução, no qual apresentamos também a seção com a justificativa dessa pesquisa.

No capítulo 2 apresentamos o referencial teórico, no qual abordamos as seguintes temáticas: pragmática e semântica, semântica na OC e na recuperação da informação (RI), tesouros, relações semânticas e normas para elaboração de tesouros.

No capítulo 3 descrevemos os procedimentos metodológicos adotados e a caracterização da pesquisa.

O capítulo 4 contém a descrição dos resultados, que foi dividida em três seções: abordagem pragmática na CI e na OC; como os elementos dessa abordagem se encontram nas normas para elaboração

de tesouros; e orientações para o estabelecimento de relações semânticas em tesouros.

Para finalizar, apresentamos as considerações finais dessa pesquisa no capítulo 5.

1.1 OBJETIVOS

Para responder a questão da pesquisa e orientar o desenvolvimento do trabalho foram definidos os seguintes objetivos:

1.1.1 Objetivo geral

Analisar como a abordagem semântica pragmática é ou pode ser adotada para o estabelecimento de relações semânticas em tesouros.

1.1.2 Objetivos específicos

- a) Caracterizar a abordagem semântica pragmática no campo de OC;
- b) Identificar a compatibilidade entre as características da abordagem pragmática e as recomendações das normas para elaboração de tesouros;
- c) Sistematizar um conjunto de orientações para o estabelecimento de relações semânticas em tesouros, segundo o paradigma pragmático.

1.2 JUSTIFICATIVA

A investigação das relações semânticas justifica-se pelo importante papel que desempenham nos tesouros e demais SOC. Mas, apesar disso, não apresentam-se tanto quanto outros aspectos desses sistemas em pesquisas da área de CI e no campo de OC. Os estudos

Têm-se centrado mais em conceitos e termos, mas o foco irá cada vez mais se deslocar para a identificação, processamento e gestão das relações, para alcançar uma maior eficácia e refinamento das técnicas em ciência da

informação. (KHOO; NA, 2006, p. 158, tradução nossa²).

O estabelecimento de relações semânticas em tesouros pode ser feito de maneiras e de acordo com pressupostos teóricos diferentes. Essas diferenças, que ocorrem na aplicação, na prática de desenvolvimento de um tesouro, ou apenas nas investigações para essas práticas/aplicações, refletem diferentes paradigmas teóricos que nem sempre são apresentados de forma explícita, mas isso não significa que não existam. Considera-se que conhecê-los, diferenciá-los e evidenciar possíveis lacunas entre a teoria e a prática pode ser útil para o aprimoramento do processo de OC, processo no qual se insere o desenvolvimento de tesouros e o estabelecimento das relações semânticas.

O estabelecimento de relações semânticas em tesouros depende dos demais processos envolvidos na elaboração desses. As normas para elaboração de tesouros apresentam as práticas já consolidadas, tanto em exemplos como nas suas recomendações.

A investigação da abordagem semântica pragmática justifica-se por ser um paradigma recente em pesquisas no campo de OC, principalmente em pesquisas empíricas. A pesquisa procura aproximar a base teórica, que sustenta a abordagem semântica pragmática, da aplicação prática, refletida nas normas de elaboração de tesouros.

² have focused on concepts and terms, but the focus will increasingly shift to the identification, processing, and management of relations to achieve greater effectiveness and refinement in information science techniques. (KHOO; NA, 2006, p. 158)

2 REFERENCIAL TEÓRICO

Nesse capítulo apresentamos o referencial teórico dos temas que envolvem a pesquisa. Por se tratar de uma pesquisa que tem como um dos objetivos analisar a abordagem pragmática na OC, tendo por base o referencial teórico encontrado na literatura da área, nessa seção apresentamos alguns temas e conceitos relacionados com o pragmatismo, a pragmática e o paradigma pragmático em semântica com o intuito de introduzir a discussão sobre o assunto que voltamos a abordar na seção 4.1 como parte dos resultados da pesquisa. Os demais temas envolvidos na pesquisa e abordados no referencial teórico são a semântica na OC e na RI, os tesauros, as relações semânticas e as normas para elaboração de tesauros.

2.1 PRAGMATISMO, PRAGMÁTICA OU PARADIGMA PRAGMÁTICO EM SEMANTICA?

Marcondes (2000) considera que é preciso distinguir a *pragmática* enquanto um campo de estudos da linguagem e o *pragmatismo* enquanto corrente filosófica, ainda que uma filosofia da linguagem na linha da pragmática e o pragmatismo se aproximem em muitos aspectos, sem contudo se confundirem. Assim, nessa seção apresentamos esta distinção que nos dá subsídios para descrever o que vem a ser o paradigma pragmático em semântica, o qual é abordado ao longo desse trabalho.

O pragmatismo, como corrente filosófica, refere-se às concepções de filosofia que defendem, sobretudo, o primado da razão prática em relação à razão teórica. Teve origem no final do século XIX e desenvolveu-se ao longo do século XX, principalmente nos Estados Unidos. “Charles Sanders Peirce (1839-1914), William James (1842-1910) e posteriormente John Dewey (1859-1952), são os principais representantes desse pensamento em suas várias vertentes.” (MARCONDES, 2000, p. 39).

A partir de uma análise dos trabalhos de Peirce, James e Dewey, considerados os fundadores do pragmatismo, Campos e Venâncio (2006) descrevem a matriz filosófica do pragmatismo:

- (1) o antifundacionalismo, que consiste em uma rejeição a qualquer espécie de entidade metafísica ou categoria apriorística, implicando uma crítica

constante; (2) o consequencialismo [sic] ou instrumentalismo, relacionado ao julgamento de proposições com base em um teste consequencialista [sic] e a uma reavaliação da verdade; (3) o contextualismo, com a valorização da experiência, da prática e dos fatos, insistindo em uma investigação compartilhada para gerar significados e ideias convergentes. Essas características pressupõem um afastamento da metafísica, a rejeição do nominalismo e uma proximidade com o realismo. Defende-se que o valor de verdade é um valor prático das ações humanas. A função do pensamento é produzir hábitos de ação, que dariam sentido às coisas. (CAMPOS; VENÂNCIO, 2006, p. 10)

Na linguística, a pragmática se caracteriza pelo estudo da linguagem em uso. Fazendo parte do campo geral de estudos da linguagem, que é consagradamente dividido entre a pragmática, a semântica e a sintaxe. No entanto, conforme esclarece Marcondes (2000), existe ainda a concepção de pragmática na qual se valoriza a linguagem comum e o uso concreto da linguagem como a principal instância de investigação da linguagem,

Segundo a qual o estudo da linguagem deve ser realizado em uma perspectiva pragmática, ou seja, enquanto prática social concreta, examinando portanto a constituição do **significado linguístico**[sic] a partir da interação entre falante e ouvinte, do contexto de uso, dos elementos sócio-culturais pressupostos pelo uso, e dos objetivos, efeitos e consequências [sic] desses usos. A pragmática não seria assim apenas um segmento dos estudos da linguagem, mas o seu campo privilegiado. (MARCONDES, 2000, p. 40 grifo nosso).

Nessa vertente o autor inclui a filosofia da linguagem ordinária de Gilbert Ryle, a teoria dos atos de fala de Austin, a concepção de jogos de linguagem de Wittgenstein, e a semiótica de Umberto Eco.

Se em suas origens o pragmatismo e a pragmática nasceram separados, verifica-se que a união de ambos gerou bons frutos, e

acreditamos que ainda pode gerar mais. Conforme descreve Marcondes (2000):

Na filosofia alemã, por sua vez, J.Habermas (*Pragmática universal* e *Teoria da ação comunicativa*) e K.O. Apel (*Pragmática transcendental*) inspirados no pragmatismo e na filosofia pragmática da linguagem, desenvolveram concepções de pragmática voltadas para a análise das condições de possibilidade da comunicação, de seus pressupostos e de suas implicações, inclusive nos campos da ética e da política. (MARCONDES, 2000, p. 40).

Surgindo o chamado neo-pragmatismo,

O foco na linguagem, uma consequência da chamada virada linguística, na filosofia ocidental, é o que, de acordo com Rorty (1999, pp. 24, 35), separa o neo-pragmatismo do pragmatismo clássico, que incide sobre a experiência. Essa mudança de foco implica que, em vez de se concentrar nas experiências de mentes individuais, há uma ênfase na experiência de comunicação destas através da linguagem. (SUDIN; JOHANNISSON, 2005, p. 25, tradução nossa³).

O pragmatismo como corrente filosófica, assim como o positivismo, a fenomenologia, entre outras, influenciam as pesquisas de diferentes áreas do conhecimento, inclusive a área de estudos da linguagem. Os pesquisadores em epistemologia definem esta influência como paradigma ou visão de mundo. Mas um paradigma não é necessariamente apenas uma influência que ocorre posteriormente ao desenvolvimento de uma corrente filosófica, mas, também pode ser um enquadramento/compatibilidade com uma determinada corrente.

Quando alguns dos praticantes da ciência descobrem contradições internas e chegam à conclusão de que a forma de ver o mundo em que

³ The focus on language, a consequence of the so-called linguistic turn in Western philosophy, is what, according to Rorty (1999, pp. 24, 35), separates his neo-pragmatism from classical pragmatism, which focuses on experience. This shift in focus implies that rather than concentrating on the experiences of individual minds, there is an emphasis on the communication of these experiences through language.

essa ciência se baseia não é adequada, começam a descobrir que o mundo deveria, ou poderia ao menos, ser olhado e investigado de outra maneira (KUHN, 2001).

Na área de estudos da linguagem encontramos ainda estudos epistemológicos que consideram a semântica como a área principal, sendo o significado o seu objeto de estudo, e a pragmática um paradigma em semântica. Hjørland (2007b) ressalta que Peregrin (2004) destaca dois paradigmas dominantes em semântica, o pragmático e o positivista. O paradigma semântico positivista seria a semântica formal, com regras de lógica para inferências, no qual a linguagem é vista como uma forma de representação do mundo. No paradigma pragmático em semântica a linguagem é vista como um meio de interação e o significado de uma expressão não é a entidade representada por ela, mas sim a sua função na interação, aceitando-se, portanto, diferentes interpretações.

Peregrin (1999) trata especificamente sobre a pragmatização da semântica e afirma que Charles Morris e Rudolf Carnap estabeleceram uma divisão triádica para as teorias da linguagem: **sintaxe** trataria as relações entre as expressões; **semântica** abordaria as relações entre as expressões e o que elas representam (significados); **pragmática** examinaria as relações entre as expressões e os seus usos. No entanto, assim como Peregrin (1999, 2004), entendemos que a semântica e a pragmática deveriam ser estudadas juntas, a semântica formal não deveria ser desprezada, mas entendida “como uma forma de destacar e materializar cada expressão contribuindo para as inferências de acordo com o contexto onde ocorrem” (PEREGRIN, 2004, p. 12, tradução nossa⁴).

Com a pragmática foi alterada e implementada a forma de estudar o significado, ainda que o pragmatismo (como corrente filosófica) possa extrapolar os estudos da linguagem e do significado, nessa pesquisa aceitamos a pragmática como um dos paradigmas em semântica. Pois, o estudo do significado tem grande relevância para a OC e RI, conforme é descrito no próximo tópico.

⁴ As a way to single out and materialize each expression's contribution to the inferences in which it occurs.

2.2 SEMÂNTICA NA ORGANIZAÇÃO DO CONHECIMENTO E NA RECUPERAÇÃO DA INFORMAÇÃO

Na literatura da área de CI é evidenciada a ligação da área de semântica com os campos de OC (ALMEIDA; SOUZA; FONSECA, 2010; CAFÉ; BRASCHER, 2011; HJØRLAND, 2007b; MAZZOCCHI, TIBERI, SANTIS, PLINI, 2007). Existe um consenso entre os diferentes autores que a área de semântica pode fornecer subsídios teóricos para o campo de OC.

Para explicar essa ligação no campo de OC, é apontado o fato de os SOC serem ferramentas semânticas. (HJØRLAND, 2007b; MAZZOCCHI; TIBERI; SANTIS; PLINI, 2007). Uma vez que fornecem um “roteiro semântico” por meio das relações que estabelecem entre os conceitos. Assim, de acordo com Hjørland (2007b):

O conhecimento sobre conceitos e semântica é importante para a investigação e utilização de qualquer um desses sistemas. Em outras palavras, pesquisadores em OC devem fundamentar o seu trabalho em uma teoria de semântica fértil. Este tipo de pesquisa básica, no entanto, é praticamente inexistente em CI. (HJØRLAND, 2007b, p. 369, tradução nossa⁵).

Na RI, Blair (2003) considera que, devido ao crescimento dramático e aparentemente inevitável no tamanho dos sistemas de recuperação da informação e as muitas maneiras que as descrições de informações podem dar errado, é importante examinar de perto a atividade de descrever o conteúdo intelectual da informação para aprimorar os processos de descrever o que se quer (necessidades de informação) e descrever o que está disponível (representação temática). É necessário, ainda, entender uma afirmação que a princípio pode parecer óbvia, “quando nós descrevemos o que queremos, essa descrição deve significar alguma coisa. Mas o que exatamente queremos dizer

⁵ knowledge about concepts and semantics is important for research into, and the use of, any of those systems. In other words, researchers in KO should ground their work in a fruitful theory of semantics. This kind of basic research has, however, been largely absent from IS.

quando descrevemos o que queremos?” (BLAIR, 2003, p. 8, tradução nossa⁶).

Van Rijsbergen (1986, apud BLAIR, 2003; HJØRLAND, 2007b) aponta que o conceito de significado é um dos componentes visivelmente mais ausentes da teoria de recuperação de informação na CI. Sobre esse aspecto, Hjørland (2007b) enfatiza a importância da semântica para a OC e Blair (2003) considera que a filosofia da linguagem pode fornecer algumas orientações, pois, como destaca:

Os filósofos têm refletido sobre o “significado do significado” desde pelo menos o tempo de Aristóteles, mas talvez nenhum filósofo teve mais impacto sobre a filosofia da linguagem do que Ludwig Wittgenstein (1889-1951).[...] Os escritos de Wittgenstein sobre a filosofia da linguagem são “extensos e intimamente ligados às suas visões sobre a filosofia da mente. Para ele, a linguagem, não é um produto do pensamento, como a maioria dos filósofos aceita; “linguagem”, como ele mesmo disse, “é... o veículo do pensamento.” (BLAIR, 2003, p. 9 e 13, tradução nossa⁷).

Conforme descrevemos, o estudo da linguagem e do significado são essenciais na fundamentação dos estudos em OC e na RI. Isso faz com que as diferentes formas de ver e estudar a linguagem e o significado gerem diferenças marcantes nas pesquisas de CI. De acordo com Campos e Venâncio (2006, p. 10) “o pragmatismo e a filosofia da linguagem tiveram papel preponderante na diversificação das abordagens no campo da Ciência da Informação”.

Esta diversificação das abordagens, bem como a influência dos estudos da linguagem nestas, pode ser identificada em estudos epistemológicos da CI. Capurro (2003), por exemplo, destaca três paradigmas na área, o físico, o cognitivo e o social. O **paradigma físico** é relacionado com a chamada *information theory* de Claude Shannon e

⁶ “when we describe what we want, we *must* mean something by that description.” But what exactly *do* we mean when we describe what we want?

⁷Philosophers have pondered the “meaning of meaning” since at least the time of Aristotle, but perhaps no philosopher has had more impact on the philosophy of language than Ludwig Wittgenstein (1889-1951). [...] Wittgenstein’s writings on the philosophy of language were extensive and closely linked to his views on the philosophy of mind. To him, language is not a *product* of thought, as most philosophers accepted; “language,” as he put it, “is . . . the vehicle of thought.”

Warren Weaver (1949-1972) e também com a cibernética de Norbert Wiener (1961). Em essência esse paradigma postula que há algo, um objeto físico, que um emissor transmite a um receptor. A ontologia e a epistemologia de Karl Popper influenciaram diretamente o **paradigma cognitivo** proposto por B. C. Brookes (1977, 1980), entre outros. Peter Ingwersen, de acordo com o autor, também é um representante desse paradigma. Em suas pesquisas, apesar de apresentar certo enfoque social,

Sua perspectiva permanece cognitiva no sentido de que se trata de ver de que forma os processos informativos transformam ou não o usuário, entendido em primeiro lugar como sujeito cognoscente possuidor de “modelos mentais” do “mundo exterior” que são transformados durante o processo informacional. Ingwersen toma elementos da teoria dos „estados cognitivos anômalos“ (“anomalous state of knowledge” abreviado: ASK), desenvolvida por Nicholas Belkin e outros (Belkin 1980, Belkin/Oddy/Brooks 1982). Essa teoria parte da premissa de que a busca de informação tem sua origem na necessidade (“need”) que surge quando existe o mencionado estado cognitivo anômalo, no qual o conhecimento ao alcance do usuário, para resolver o problema, não é suficiente. Tal situação inicial geralmente também se denomina „situação problemática“ (Wersig 1979). A teoria dos modelos mentais tem tido impacto no estudo e na concepção de sistemas de recuperação da informação, como mostram as análises empíricas realizadas por Pertti Vakkari com relação à conexão entre estado anômalo do conhecimento e estratégias de busca (Vakkari 2003). Nesse sentido, podemos falar, tanto no caso de Ingwersen quanto no de Vakkari, de uma posição intermediária entre o paradigma cognitivo mentalista de Brookes e o paradigma social. (CAPURRO, 2003).

O ponto de vista cognitivo, no entanto, é considerado idealista e também associial por Bernd Frohmann. “A crítica de Frohmann é baseada em parte implicitamente na epistemologia do Wittgenstein das

“Investigações Filosóficas” (Wittgenstein 1958), bem como na teoria do discurso como manifestação de poder, de Michel Foucault (1994)”. (CAPURRO, 2003). A partir desta crítica o autor descreve o **paradigma social**, apresentando como principal exemplo deste a análise de domínio desenvolvida por Birger Hjørland e Hanne Albrechtsen.

Uma consequência prática desse paradigma é o abandono da busca de uma linguagem ideal para representar o conhecimento ou de um algoritmo ideal para modelar a recuperação da informação a que aspiram o paradigma físico e o cognitivo. Uma base de dados bibliográfica ou de textos completos tem caráter eminentemente polissêmico ou, como o poderíamos chamar também, polifônico. Os termos de um léxico não são algo definitivamente fixo. O objeto da ciência da informação é o estudo das relações entre os discursos, áreas de conhecimento e documentos em relação às possíveis perspectivas ou pontos de acesso de distintas comunidades de usuários (Hjørland 2003). Isso significa, em outras palavras, uma integração da perspectiva individualista e isolacionista do paradigma cognitivo dentro de um contexto social no qual diferentes comunidades desenvolvem seus critérios de seleção e relevância (CAPURRO, 2003, p. 12).

Além dos paradigmas inerentes à área de CI, que são o físico, o cognitivo e o social, conforme definição de Capurro (2003). Hjørland (2003) identifica a influência de correntes epistemológicas mais amplas que também influenciam o campo de OC na CI. Estas correntes, não deixam de ser uma outra classificação dos paradigmas. Só que nesse caso baseado em estudos epistemológicos gerais aplicados para a classificação dos paradigmas na OC especificamente, e não na CI como um todo.

Hjørland (2003), considera que os métodos de OC, na CI, estão relacionados com pressupostos epistemológicos, dos quais o autor destaca o empirismo, o racionalismo, o historicismo e o pragmatismo. Segundo o autor os pesquisadores em qualquer campo sempre são mais ou menos influenciados por certas ideais sobre como obter conhecimento. Se estudantes de CI fizessem um curso de filosofia da ciência, poderiam ser capazes de identificar a abordagem epistêmica

dominante em qualquer artigo sobre a OC e também em qualquer SOC. Alguns artigos e sistemas são baseados principalmente em generalizações empíricas, como por exemplo os sistemas com base na medida de frequência de palavras. Outros artigos e sistemas são na sua maioria baseado em regras racionais e deduções (embora muitas vezes ignorando questões empíricas), neste caso o autor considera que os sistemas analítico-facetados na tradição de Ranganathan podem fornecer os melhores exemplos. Um terceiro tipo de sistema é baseado no estudo da evolução de campos do conhecimento (historicismo). Em certo grau os sistemas com bases disciplinares, como a Classificação Decimal de Dewey (CDD) estariam em conformidade com este ideal. Mas, por outro lado, a atualização deste sistema é mais influenciada pelos princípios da tradição analítico-facetada (como discutido por Miksa, 1998). E, em certo sentido, a CDD também tem base empírica. Assim, o autor considera que no caso da CDD os seus princípios não são tão claros quanto nos exemplos anteriores. No caso dos estudos e sistemas com base no pragmatismo, considera-se que estes são desenvolvidos com base na análise dos objetivos, valores e consequências. Por exemplo, pesquisadoras mulheres podem considerar a meta de emancipação das mulheres como o objetivo principal dos estudos feministas. Elas podem, portanto, considerar quaisquer relações que suportam esta meta importante. “Similaridade” é, portanto, identificada por classes de documentos compartilhando funções similares em relação a este objetivo. Classificação pragmática pode assim também ser denominada classificação crítica ou classificação política. O que poderia parecer à primeira vista estranho e oposto as normas científicas comuns.

Mas, conforme o autor esclarece,

A epistemologia pragmática na OC não significa que uma pessoa (ou um campo inteiro) pode simplesmente fazer as coisas a maneira que se adapte a seus interesses pessoais (ou os interesses dos pesquisadores da área). Se isso for feito, se a pesquisa apenas produz “construções sociais”, em seguida, a realidade vai mostrar que essas construções são incoerentes. Elas vão ser contestadas por argumentos teóricos e empíricos. A produção de “conhecimento” incoerente não tem valor e não pode ser um objetivo sério. O método pragmático não se opõe aos aspectos do empirismo, racionalismo e historicismo. Alega, no entanto, que evidências isoladas não são

suficientes. Os critérios finais de verdade estão ligados aos objetivos e às atividades humanas. Não se pode deixar de considerar essas questões, embora elas possam parecer desconfortáveis. (HJØRLAND, 2003, p. 106, tradução nossa⁸).

As classes de paradigmas de Capurro (2003) e Hjørland (2003) não são mutuamente excludentes, na nossa opinião. Assim, podemos dizer que o paradigma pragmático na OC é compatível com o paradigma social na CI.

2.3 TESAUROS

O tesauro surgiu da necessidade de manipular grande quantidade de documentos especializados, o que exigia trabalhar com um vocabulário mais específico e com uma estrutura mais depurada do que aquela presente nos cabeçalhos de assunto (remissivas e referências cruzadas tipo ver e ver também). Assim, além da especificidade, cuidou-se de melhorar a estrutura e as referências cruzadas, que deram lugar às relações semânticas. Pelo fato desse novo instrumento da documentação possibilitar, por meio do agrupamento dos termos, o acesso a uma ideia, mesmo sem saber nomeá-la de saída, as novas listas estruturadas de termos passaram a ser chamadas de tesauros, por analogia com o termo “tesauro” que tem origem no dicionário analógico de Peter Mark Roget, intitulado *“Thesaurus of English words and phrases”*, publicado, pela primeira vez, em Londres, em 1852. (GOMES, 1990).

Gomes (1990) esclarece que Roget era Secretário da *Royal Society*⁹ e pretendia, com aquela obra, facilitar sua atividade literária. Em seu dicionário as palavras não foram agrupadas segundo a ordem

⁸ Pragmatic epistemology and pragmatic KO do not mean that a person (or a whole field) can just do things the way that suits his personal interests (or the interests of the researchers in the field). If this is done, if research just produces “social constructions” then reality will make those constructions incoherent. They will be opposed by empirical and theoretical arguments. The production of incoherent “knowledge” is not valuable and cannot be a serious goal. Therefore pragmatic philosophy is bound up with a form of realism. The pragmatic method is not opposed to aspects of empiricism, rationalism and historicism. It claims, however, that no isolated evidence is enough. The final criteria of truth are connected to human goals and activities. You cannot avoid considering such issues although they may seem uncomfortable.

⁹ É uma instituição destinada à promoção do conhecimento científico, fundada em 28 de novembro de 1660.

alfabética, como ocorre com os dicionários da língua, mas “de acordo com as ideias que elas exprimem”. Conforme explica:

O propósito de um dicionário comum é simplesmente explicar o significado das palavras; e o problema para o qual ele pretende oferecer a solução pode ser apresentado assim: - Sendo dada uma palavra, encontrar seu significado ou a ideia que ela pretende trazer consigo. O que se almeja com este empreendimento é exatamente o contrário: a saber, - Tendo-se a ideia, encontrar a palavra, ou as palavras, pelas quais a ideia possa ser expressa de maneira mais adequada e ajustada. Com este objetivo, as palavras e frases da língua estão arranjadas aqui não de acordo com seu som ou sua ortografia, mas estritamente de acordo com seu significado. (ROGET, apud GOMES, 1990, p.13)

Com este objetivo, dentre outros, e seguindo este tipo de arranjo também se desenvolve o tesouro documentário, que é definido por Currás (1995, p. 88) como “uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos linguísticos que o compõem – termos simples ou compostos – encontram-se relacionados entre si sintática e semanticamente”.

Bräscher (2010) apresenta um complemento para esta definição, esclarecendo que o tesouro é, essencialmente, um sistema pós-coordenado, apesar de possibilitar maior ou menor nível de pré-coordenação de descritores. E pode ser definido,

Quanto à função, os tesouros são instrumentos de controle terminológico utilizados em sistemas de informação para traduzir a linguagem dos documentos, dos indexadores e dos pesquisadores numa linguagem controlada, usada na indexação e na recuperação de informações.

Quanto à estrutura, o tesouro é um vocabulário controlado e dinâmico de termos (descritores) semanticamente relacionados, que cobrem uma determinada área do conhecimento. (BRASCHER, 2010).

A estrutura de um tesauro pode variar e diferentes alternativas podem ser empregadas para o controle da linguagem. Mas um tesauro é composto basicamente por termos (preferidos e não-preferidos), qualificadores e relações semânticas. Os termos representam os conceitos da área do conhecimento que o tesauro cobre. Os qualificadores são palavras ou expressões que normalmente aparecem entre parênteses junto de alguns termos, com a função de eliminar ambiguidades de palavras homógrafas ou polissêmicas. As relações empregadas nos tesouros são denominadas como: relações de equivalência (entre os termos preferidos e não-preferidos), relações hierárquicas (entre termos gerais e específicos) e relações associativas (ligações entre os termos que não são equivalentes e nem hierárquicos, mas que ainda assim são termos semanticamente ou conceitualmente associados a um ponto tal que se considera necessário explicitar a ligação entre eles).

Diferentemente das ontologias e redes semânticas, que tem origem no contexto da era digital, os tesouros tiveram origem no contexto das bibliotecas tradicionais. Foram instrumentos criados, originalmente, para auxiliar os bibliotecários na indexação. Posteriormente, principalmente quando inseridos no contexto digital, passaram a ser disponibilizados também aos usuários, para auxiliar a busca por informações.

A maior parte dos tesouros estabelecem um vocabulário controlado, padronizando a terminologia, no qual cada conceito é representado por um termo, um descritor, que é usado na indexação e pode assim, ser utilizado com confiança na busca. O tesauro deve oferecer suporte ao indexador para identificar todos os descritores que devem ser atribuídos a um documento à luz das questões que possam vir a ser feitas. Um bom tesauro fornece, através de sua organização hierárquica, estendida por relações associativas entre conceitos, um roteiro semântico para pesquisadores e indexadores e qualquer pessoa interessada em uma compreensão ordenada de um campo de assunto. (SOERGEL, 1995, p.369, tradução nossa)

A organização básica dos tesouros é hierárquica, ou seja, organizada do geral para o específico, formando classes/categorias. Para

a formação dessas classes existem dois métodos comumente empregados, o método indutivo (bottom-up) e o método dedutivo (top-down). Batista (2004, p. 9) relaciona o princípio de garantia literária¹⁰ ao método indutivo na OC e afirma que, de acordo com este princípio, “a determinação de classes deve ser feita conforme revelada na literatura, e não a partir de qualquer classificação do conhecimento preexistente”. O autor também sustenta que:

O método dedutivo tende a revelar a estrutura paradigmática dominante de uma área do conhecimento, por basear-se no consenso entre pares, mas pode excluir paradigmas emergentes. O método indutivo - via literatura corrente - pode detectar **novas relações** entre conceitos. (BATISTA, 2004, p. 9, grifo nosso).

Entretanto, estas novas relações não se referem às relações hierárquicas, apesar da afirmação anterior – sendo que a formação de classes se dá por meio destas –, mas sim às relações associativas, não especificadas, o que fica evidenciado no decorrer do artigo do autor supracitado. No entanto, destacamos que identificar na literatura os conceitos que possivelmente integrem uma relação não é identificar a relação em si.

Considerando que as relações semânticas são elementos essenciais para dar consistência e coerência conceitual a um tesouro, Barité et al. (2010, p. 134) afirmam que, “deve-se perguntar se a garantia literária não é capaz de gerenciar novas técnicas, para garantir a pertinência das relações presentes nos sistemas de organização do conhecimento (de sinonímia, hierárquicas ou associativas)”. Os autores também afirmam que:

Existe, pois, um amplo espaço a ser explorado para determinar se a garantia literária pode ser utilizada como uma ferramenta útil para a distinção entre relações paradigmáticas e sintagmáticas, e para legitimar aquelas relações paradigmáticas que, por sua fortaleza e

¹⁰ A concepção original de garantia literária se sustenta na ideia nuclear de que a literatura de um domínio deve ser a fonte para extração e validação da terminologia a ser incorporada em um sistema de classificação, ou em qualquer outro sistema de organização do conhecimento. (BARITÉ; FERNÁNDEZ-MOLINA; GUIMARÃES; MORAES, 2010, p. 124)

estabilidade, deveriam estabelecer-se em um tesouro ou em outro vocabulário, sob a forma de relações hierárquicas ou associativas. (BARITÉ; et al. 2010, p. 134).

Tendo em vista o papel fundamental das relações semânticas em SOC e, mais especificamente, em tesouros, conforme explicamos, passaremos na próxima seção ao aprofundamento dessas relações utilizadas em tesouros, assim como as relações semânticas em geral.

2.4 RELAÇÕES SEMÂNTICAS

Relações semânticas existem independentemente de serem utilizadas/explicitadas em um SOC. De acordo com Khoo e Na (2006, p.159, tradução nossa¹¹), “são associações significativas entre dois ou mais conceitos, entidades ou grupos de entidades. Elas podem ser vistas como ligações direcionais entre os conceitos/entidades que participam da relação”. Estes conceitos e entidades não estão presentes apenas em SOC.

Conceitos são como os blocos de construção do conhecimento, relações funcionam como o cimento que une conceitos em estruturas de conhecimento (KHOO; NA, 2006; GREEN; BEAN; MYAENG, 2002). Nesta união, a relação, se dá por meio das características dos conceitos.

Sempre que diferentes conceitos possuem características idênticas deve-se admitir que entre eles existem relações. Este fato tem importância na ordenação dos conceitos. Se tivermos os seguintes conceitos: *semear, colher, transportar a colheita, preparar o solo, arar a solo*, existe uma característica comum que é *atividade agrícola*. Então ha de convir que tais conceitos possuem relações entre si. (DAHLBERG, 1978, p. 104).

Os “blocos da construção”, os conceitos, são definidos por Dahlberg (1978, p. 102) como “a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico”. Para a autora, “todo enunciado sobre objetos contem um elemento do

¹¹ are meaningful associations between two or more concepts, entities, or sets of entities. They can be viewed as directional links between the concepts/entities that participate in the relation.

respectivo conceito. Estes elementos identificam-se com as chamadas *características dos conceitos*”. Khoo e Na (2006, p. 159) esclarecem que “os conceitos e entidades são parte integrante da relação, uma vez que uma relação não pode existir apenas por si mesma”.

As relações podem ser categorizadas por razões e de maneiras diferentes. Dahlberg (1978) as categoriza de acordo com a posse, ou não, de características comuns, afirmando que, de acordo com este princípio, as seguintes relações são logicamente possíveis: identidade, implicação, intersecção, disjunção e negação. Khoo e Na (2006), esclarecem que:

Associações entre os conceitos/entidades podem ser categorizadas em diferentes tipos – abstração, conceitualização e distinção, entre outros – e podem assim ter significado atribuído. O significado ou tipo de associação pode, por vezes, mas não sempre, ser derivado do significado dos conceitos envolvidos. Psicólogos e filósofos¹² têm procurado identificar os principais tipos de relações e suas características. (KHOO; NA, 2006, p. 159, tradução nossa¹³)

Para Hjørland (2007b), relações entre conceitos não devem ser confundidas com as relações entre os termos que são usados para expressar os conceitos. É comum no entanto, reunir esses dois tipos de relações sob o título “relações semânticas”. Khoo e Na (2006, p. 160, tradução nossa¹⁴) também afirmam que “muitas vezes a distinção entre relações lexicais e conceituais são irrelevantes e autores usam o termo relações léxico-semânticas”.

Mesmo não existindo um total consenso entre os estudiosos, com relação à estas e outras questões, Khoo e Na (2006) esclarecem que,

¹² Incluímos ainda os linguistas ao grupo de profissionais que investigam as relações semânticas. Além, é claro, dos profissionais do campo de OC, que atualmente têm direcionado seus esforços ao estudo das relações semânticas. Ainda que no caso destes últimos, o estudo das relações semânticas se refira quase que exclusivamente ao uso destas em SOC's, e não à investigação de relações semânticas em si.

¹³ Associations between concepts/entities can be categorized into different types, abstracted, conceptualized and distinguished from other associations, and can thus be assigned meaning. The meaning or type of an association can sometimes but not always be derived from the meanings of the concepts involved. Psychologists and philosophers have attempted to identify the main types of relations and their features.

¹⁴ Often the distinction between conceptual relations and lexical relations are unimportant and authors use the term *lexical-semantic relation*.

sobre a classificação dos tipos de relações, a maioria dos pesquisadores reconhecem duas grandes categorias: as paradigmáticas e as sintagmáticas. Asher e Simpson (1994, p. 5178) afirmam que as relações sintagmáticas se referem à “relação que uma unidade linguística tem com as outras unidades com as quais co-ocorre numa sequência ou contexto.” Lancaster (1986) caracterizou as relações sintagmáticas como *a posteriori* ou relações transitórias/ocasionais. De acordo com Barité et al. (2010, p. 134):

As relações ocasionais ou eventuais entre conceitos que se expressam na documentação, também chamadas de sintagmáticas (Gardin *et al.*, 1968; Tálamo, 1997), não precisam ser transladadas para o sistema de organização do conhecimento, sempre que, em seu lugar, as linguagens pré-coordenadas ofereçam mecanismos ou sinais para sua conexão (por exemplo, sinais como “:” ou “+” na Classificação Decimal Universal).

As relações paradigmáticas, seriam “um conjunto de relações entre as unidades linguísticas, onde uma unidade pode substituir a outra de acordo com diferentes ambientes linguísticos” (ASHER; SIMPSON, 1994, v. 10, p. 5153). Por exemplo, nas frases: “Eu comi uma laranja / A laranja foi comida por mim”. Eu e mim são substitutos nestes dois contextos, e a relação que existe entre estes dois termos é paradigmática. (ASHER; SIMPSON, 1994). Lancaster (1986) caracterizou relações paradigmáticas como *a priori* ou relações permanentes. Barité et al. (2010, p. 134) afirmam que as relações paradigmáticas “são permanentes em função de haver nexos ou vínculos estáveis entre os objetos aos quais se referem, devido a sua origem, suas propriedades ou características”.

No entanto, Khoo e Na (2006, p. 166, tradução nossa¹⁵) esclarecem que “a distinção entre relações paradigmáticas e sintagmáticas é imprecisa. [...] as relações paradigmáticas podem ser expressas no nível sintático”. Os autores procuram reforçar a explicação desta distinção com a seguinte afirmação de Evens (1980):

¹⁵ The distinction between paradigmatic and syntagmatic relations is fuzzy. [...] paradigmatic relations can be expressed syntagmatically.

Parece que estamos recebendo informação paradigmática normalmente em sentenças genéricas (sempre verdadeiras), enquanto que relações sintagmáticas vêm a nós em orações ocasionais. Uma sentença genérica ou permanente contém alguma informação permanente sobre o mundo, como ‘O alimento é comestível’. Uma frase ocasional contém informações sobre um contexto particular. (EVENS, 1980, apud KHOO; NA 2006, p. 166, tradução nossa¹⁶).

Considerando que, “as principais relações léxico-semânticas são as relações paradigmáticas de hponímia (gênero-espécie), relação todo-parte, sinonímia, e antonímia” (KHOO; NA, 2006, p. 167), pode-se inferir que os conceitos e as principais relações semânticas que compõem os SOC baseiam-se em sentenças ‘sempre verdadeiras’, que são menos comuns do que as ocasionais. As sentenças ‘sempre verdadeiras’ podem ser insuficientes para a implementação dos componentes de um SOC, pois um determinado domínio do conhecimento pode apresentar pontos de vista divergentes, o que traz a questão relativa ao desenvolvimento desses sistemas de forma consistente¹⁷.

Nesse sentido, também existem duas grandes abordagens para a construção de uma lista de relações semânticas que são, de acordo com Evens (1988 apud KHOO; NA, 2006, p. 165) desenvolvidas, na área de semântica, por dois grupos distintos de pesquisadores, os “*lumpers*” e “*splitters*”.

Os *lumpers*, ou minimalistas, definem um pequeno número de relações gerais com base em princípios filosóficos ou lógicos (por exemplo, Sowa, 1984, 2000; Werner, 1988). Werner (1988) usa apenas três relações: *modification*, *taxonomy*, e *queuing*. O outro grupo de pesquisadores têm

¹⁶ We seem to receive paradigmatic information typically in generic (always true) sentences, while syntagmatic relationships come to us in occasional sentences. A generic or standing sentence contains a piece of permanent information about the world, such as ‘Food is edible.’ An occasional sentence contains information about a particular context. (Evens et al., 1980, pp. 10-11)

¹⁷ Uma vez que as relações paradigmáticas são insuficientes mas as sintagmáticas podem ser infinitas, pois depende do contexto, qual critério pode ser usado para definir qual relação sintagmática será considerada? Ao longo dessa pesquisa procuramos responder essa questão.

uma lista muito mais elaborada de relações específicas, muitas vezes com base em relações léxico-semânticas e palavras encontradas em um texto (por exemplo, Calzolari, 1988). (KHOO; NA 2006, p. 165)

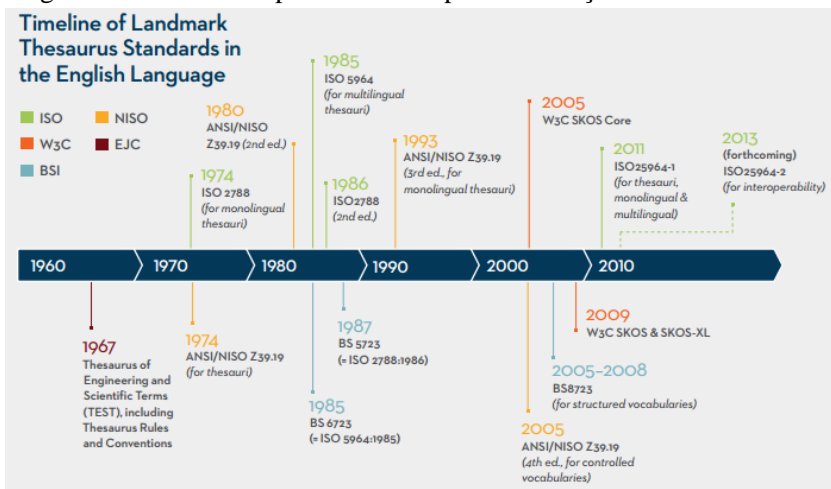
As relações definidas pelos *splitters* não são baseadas em princípios lógicos aplicados a sentenças verdadeiras (características dos conceitos), mas sim na análise de ocorrências de palavras e conceitos em textos. Firth (1957, p 195) afirma que, “você deve conhecer uma palavra pela companhia que esta mantém”. Dessa forma, infere-se que as relações definidas pelos *lumpers* seriam apenas as paradigmáticas. E as relações definidas pelos *splitters* poderiam ser tanto paradigmáticas quanto sintagmáticas.

As pesquisas sobre as relações semânticas como podemos notar ocorrem em diferentes campos de estudo, como na filosofia e na linguística. Na OC, como dissemos anteriormente, as relações semânticas são pesquisadas por que estão presentes nos tesauros, e demais SOC, como elementos estruturantes que representam as relações que ocorrem entre as entidades do mundo real. São portanto, contempladas nas normas de elaboração de tesauros apresentadas na seção a seguir.

2.5 NORMAS PARA ELABORAÇÃO DE TESAUROS

As normas para a elaboração de tesauros sofreram transformações ao longo do tempo e, assim como os tesauros, também evoluíram e se adaptaram ao desenvolvimento tecnológico e teórico da área de CI. Clarke e Zeng (2012) apresentam uma linha do tempo com os marcos das transformações ocorridas, que também correspondem à data de publicação de novas normas ou de atualizações das já existentes.

Figura 1: Linha do tempo das normas para elaboração de tesouros



Fonte: Clarke e Zeng (2012, p. 22).

Na década de 1960 houve a publicação de número significativo de tesouros bem como guias para a sua elaboração. O mais influente foi o *Thesaurus of Engineering and Scientific Terms (TEST)* em 1967, com regras e convenções para tesouros em seu apêndice. Dentre as convenções que hoje ainda prevalecem está o uso das abreviações BT, NT e RT¹⁸ para identificar relações entre conceitos (CLARKE; ZENG, 2012).

Na década seguinte, a de 1970, a primeira edição da norma americana, elaborada pela *National Information Standards Organization (NISO)* que pertence ao *American National Standards Institute (ANSI)*, e da norma internacional, elaborada pela *International Organization for Standardization (ISO)*, foram publicadas.

A norma americana, ANSI/NISO Z39.19, da primeira até a terceira edição em 2003 abrangia apenas os tesouros. Na sua quarta e última edição em 2005 ela passou a tratar sobre demais tipos de vocabulários controlados, alterando inclusive seu título (de *Guidelines for the construction, format and management of monolingual Thesauri* para *Guidelines for the construction, format and management of monolingual controlled vocabularies*). Dos quatro diferentes tipos de

¹⁸ Em português BT corresponde a TG (termo geral), NT a TE (termo específico) e RT a TR (termo relacionado) ou TA (termo associado).

vocabulários controlados que a norma trata – listas de assuntos, anéis de sinônimos, taxonomias e tesauros – o mais complexo é o tesauro.

Sobre as atualizações significativas ocorridas na Z39.19-2005, destacam-se, inicialmente, a incorporação de uma outra finalidade às quatro definidas pela norma Z39.19- 2003 para o controle de vocabulários, correspondente ao item “Etiqueta e Navegação”, enfatizando ainda mais a construção de vocabulários disponíveis eletronicamente:

- 1) Tradução: prover um modo para traduzir a linguagem natural dos autores, indexadores e usuários para um vocabulário controlado usado para indexação e recuperação;
- 2) Consistência: promover consistência na designação de termos de indexação;
- 3) Indicação de relacionamentos: indicar relacionamentos semânticos entre termos;
- 4) Etiqueta e navegação: provê hierarquias claras e consistentes em um sistema de navegação para auxiliar usuários na localização de documentos desejados;
- 5) Recuperação: servir como uma ajuda na busca e recuperação de documentos. (BOCCATO, 2011, p.171)

A norma apresenta um conjunto de recomendações com base em técnicas e procedimentos preferenciais e opcionais, utilizando as seguintes convenções para indicar a relevância das recomendações: obrigatório, recomendado, permitido, não recomendado e não permitido, e estabelece quatro princípios para a construção de vocabulários controlados:

- a) a eliminação de ambiguidades : causadas por homógrafos e polissemias;
- b) o controle de sinônimos: sinônimos, quase-sinônimos e variantes lexicais;
- c) o estabelecimento das relações lógico-semânticas entre os termos;
- d) a aplicação de testes para a validação dos termos (ANSI/NISO Z39.19:2005).

A norma internacional ISO2788 teve sua primeira edição publicada em 1974 e a sua mais recente edição foi em 1986, sendo substituída em 2011 pela ISO 25964-1, que em 2013 foi complementada pela 25964-2. De acordo com Clarke e Zeng (2012) a intenção da ISO 2788 era fornecer diretrizes para representar conceitos sem ambiguidades por meio de termos. No entanto, não havia nenhum modelo de dados explícito e a diferença entre termos e conceitos não foi articulada claramente.

Em 1985 foi publicada a ISO 5964, baseada na ISO 2788 e orientada ao contexto de tesouros multilíngues. Clarke e Zeng (2012) afirmam que esta norma sofreu com a mesma falta de clareza na distinção entre termos e conceitos.

A norma britânica BS8723 (*British Standard 8723*):

Teve sua trajetória marcada pelo desenvolvimento, inicialmente de duas edições que se configuraram, respectivamente, nas normas BS 5723, publicadas em abril de 1979 e posteriormente em setembro de 1987, ambas as edições denominadas por *Guidelines for the establishment and development of monolingual thesauri*. Essas normas (BS 5723-1979, BS 5723-1987) tiveram como correspondentes a norma ISO 2788-1986. (BOCCATO, 2011, p. 172)

O foco da BS 5723 é a construção de tesouros monolíngues, com destaque para os relacionamentos conceituais, simbologias e abreviaturas empregadas, formas de apresentação e gestão. No contexto de construção dos tesouros multilíngues, encontra-se a BS 6723, equivalendo-se à ISO 5964, ambas publicadas no mesmo ano de 1985. (BOCCATO, 2011). De acordo com Clarke e Zeng (2012) as normas britânicas, BS 5723 e BS 6723, em suas últimas edições são idênticas às normas da ISO2788/1986 e ISO5964-1985, respectivamente.

De acordo com Boccato (2011, p. 173):

Tendo em vista os diversos sistemas de organização do conhecimento disponíveis na atualidade, as normas BS 5723-1987 e 6723-1985 foram unificadas e ampliadas para outras formas e estruturas de organização do conhecimento.

Sendo assim substituídas pela norma BS8723 - *Structured vocabularies for information retrieval – guide*.

A norma ISO 25964 foi baseada na BS8723 e é dividida em duas partes. A primeira parte, a ISO25964-1: *Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval*, trata dos aspectos lexicais, monolíngues e multilíngues, incluindo um modelo de dados, formatos e protocolos para intercâmbio de dados. Além de orientações sobre a construção e gestão de tesauros; traz diretrizes sobre a aplicação da análise facetada na construção de tesauros e especificações sobre a funcionalidade de *softwares* para a gestão. A segunda parte, ISO25964-2: *Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies*, aborda a interoperabilidade entre tesauros e outros vocabulários controlados. Além de trazer orientações sobre a prática de mapeamento e de arquitetura de dados.

Devido às atualizações e alterações descritas, as normas vigentes atualmente são a ISO25964-1(2011), ISO25964-2(2013), BS8723 e ANSI/NISO Z39.19:2005.

3 PROCEDIMENTOS METODOLÓGICOS

Para o alcance do objetivo geral – analisar como a abordagem semântica pragmática é ou pode ser adotada para o estabelecimento de relações semânticas em tesouros – faz-se necessário o alcance dos objetivos específicos, descritos a seguir, juntamente com as técnicas empregadas para alcançá-los.

De acordo com a abordagem do problema, a pesquisa é qualitativa, pois está pautada em análises e interpretações de conteúdos para responder uma questão. Possuindo caráter exploratório, segundo Gil (2002), estas pesquisas proporcionam maior familiaridade com o problema, com vistas a torná-lo mais explícito, tendo como objetivo principal o aprimoramento de ideias ou a descoberta de intuições. Dado o corpus de análise, a pesquisa é bibliográfica. A técnica de pesquisa empregada foi análise de conteúdo de Bardin.

Para caracterizar a abordagem semântica pragmática no campo de OC, que corresponde ao objetivo A, foi realizada uma revisão da literatura. O levantamento bibliográfico para a revisão foi realizado da seguinte forma nas seguintes fontes:

Quadro 1 - Levantamento bibliográfico

Fonte	Estratégia de busca (Campo e termo)		Documentos recuperados	Documentos selecionados
BRAPCI	Geral - todos	Pragmática	28	12
	Geral - todos	Pragmatismo	9	4
LISA	Assunto principal	<u>pragmatic</u> OR <u>pragmatism</u>	29	11
ARIST	Verificação dos sumários		2	2
Repositórios ¹⁹	<u>pragmática</u> OR <u>pragmatismo</u>		6	3 teses 3 dissertações
Web of Science	<u>Hjørland</u>		24	24
	<u>Frohmann</u>		6	6
Plataforma Lattes	Carlos Candido Almeida		8	8
	Gustavo Silva Saldanha		1	1
	Luciana de Souza Gracioso		1	1
	Maria Nélida Gonzalez de Gomez		4	4
Total 79 documentos				

Fonte: Elaborado pela autora (2014)

A busca de trabalhos dos autores listados, na *Web of Science* e na Plataforma Lattes, se deve ao fato de os mesmos estudarem a abordagem pragmática em Ciência da Informação e/ou na OC. A busca por autores da área que estudam a abordagem pragmática não foi exaustiva, e a escolha dos autores listados no quadro foi feita a partir da análise inicial dos trabalhos encontrados na LISA e na BRAPCI. A busca nas bases de dados foi atemporal, ou seja não foi feita delimitação de período, mas foram selecionados apenas os trabalhos em foi possível ter acesso ao texto completo via internet e/ou COMUT.

Os documentos selecionados (as referências dos mesmos constam no Apêndice A) foram analisados com o intuito de identificar ideias comuns ou complementares que caracterizam o paradigma pragmático. Essas ideias foram agrupadas, criando-se assim categorias. Essa análise foi realizada em três fases, definidas por Bardin (1979), pré-análise, exploração do material e tratamento dos resultados (inferência e interpretação). Quanto ao tipo de análise empregada, a categorial, a autora descreve que:

No conjunto das técnicas da análise de conteúdo, a análise por categorias é de citar em primeiro lugar: cronologicamente é a mais antiga; na prática é a mais utilizada. Funciona por operação de desmembramento do texto em unidades, em categorias segundo reagrupamentos analógicos. Entre as diferentes possibilidades de categorização, a investigação dos temas, ou análise temática, é rápida e eficaz na condição de se aplicar a discursos diretos (significações manifestas) e simples.(BARDIN, 1979, p.153).

Para Bardin (1979, p. 117):

A categorização é uma operação de classificação de elementos constitutivos de um conjunto, por diferenciação e, seguidamente, por reagrupamento segundo o gênero (analogia), com os critérios previamente definidos.

Esses critérios podem ser semânticos (categorias temáticas) e sintáticos (os verbos, os adjetivos), dentre outros. Além disso, a categorização pode empregar dois tipos de processos, que são tidos como inversos. Esses processos recebem os nomes de “por caixas”, no

qual é fornecido o sistema de categorias e repartem-se os elementos da melhor maneira possível, à medida que vão sendo encontrados, e “por milha”, no qual o sistema de categorias não é fornecido antes resultando da classificação analógica e progressiva dos elementos. O título conceitual de cada categoria somente é definido no final da operação (BARDIN, 1979).

Para o alcance do objetivo A de nossa pesquisa o critério empregado para a formação das categorias foi o semântico. Os temas foram identificados na medida em que os trabalhos que compõem o corpus da pesquisa foram analisados. Ou seja, as categorias ou classes não foram previamente definidas. O processo empregado então foi o “por milha”, como denominado por Bardin (1979). Após a categorização, foram propostos sete enunciados que, em nossa análise, sintetizam as principais características do paradigma pragmático nos estudos da OC, e nomeiam as sub-seções da seção 4.1 que discute e apresenta os resultados.

A partir do objetivo A construímos um quadro de referência (Apêndice B), que foi empregado para *identificar a compatibilidade entre as características da abordagem pragmática e as recomendações das normas para elaboração de tesouros*. As normas analisadas são a ANSI/NISO Z39.19-2005, ISO 25964-1(2011) e a ISO 25964-2(2013), que são as normas mais atuais e que vieram substituir normas anteriores. A norma BS8723 não foi analisada, apesar de ser uma das normas vigentes, devido à dificuldade de acesso e por que a norma internacional, a ISO 25964, foi elaborada com base naquela. Clarke e Zeng (2012) consideram o conteúdo das mesmas praticamente idêntico.

Para o alcance do objetivo B o critério empregado para a análise categorial foi novamente o semântico. Nesse caso as categorias, compatibilidade ou incompatibilidade com os apontamentos da abordagem pragmática, foram definidas antes da análise das normas. O processo empregado então foi o “por caixas”, como denominado por Bardin (1979). Os resultados obtidos são descritos na seção 4.2 e sistematizados em um quadro no Apêndice B.

Como resultado da análise de conteúdo caracterizamos a abordagem pragmática nos estudos da OC e verificamos como essa abordagem se encontra nas normas de elaboração de tesouros. Com base nesses resultados procuramos *sistematizar um conjunto de orientações para o estabelecimento de relações semânticas em tesouros, segundo o paradigma pragmático*. Essas orientações são descritas na seção 4.3.

4 RESULTADOS

4.1 ABORDAGEM PRAGMÁTICA

A apresentação das características da abordagem pragmática na CI e OC foi dividida em temas/tópicos. Com os títulos de cada tópico procuramos sintetizar o tema dos mesmos. As discussões apresentadas foram construídas de modo a argumentar e justificar a forma como o agrupamento das idéias foi feito, em certos momentos identificando os pontos em comum e em outros os pontos que se complementam. Ou seja, apresentamos pontos de interseção e diálogos entre autores, de modo a evidenciar os temas centrais da abordagem pragmática na CI e OC.

Cada enunciado foi formulado baseado nas ideias encontradas no *corpus* de pesquisa. Em muitos casos as afirmações dos próprios autores foram utilizadas, ou seja, citações literais, para minimizar o viés da interpretação.

4.1.1 A informação é um fenômeno construído

Frohmann (2004) apresenta o ensaio de Nunberg (1996), como exemplo de uma maneira útil de pensar sobre a informação, que desloca a atenção de questões sobre o que a informação é em si, como um campo teórico, em direção a questões de como o fenômeno da informação é construído.

Redón Rojas (1996) considera que a informação é construída e não simplesmente inventada (subjetiva) ou descoberta (objetiva). A informação não é uma simples invenção por que para sua elaboração se toma certa matéria prima (elementos objetivos que se encontram nas coisas) e tampouco algo pronto e objetivo que só precisa ser descoberto, pois é necessária a ação do homem. Os homens transformam e são transformados pelo mundo informacional em um processo de evolução constante. O conceito de informação pragmática, onde a dialética do sujeito com o mundo que o rodeia é levada em consideração deveria ser um dos pilares dos estudos em biblioteconomia, segundo o autor.

Esclarecemos que a ideia de informação como um fenômeno construído, o produto de uma síntese de elementos objetivos e subjetivos, de propriedades reais dos objetos com a atividade do homem, diz respeito à informação na CI. Uma vez que a palavra informação é usada em diferentes contextos com significados distintos,

e um desses significados pode dizer respeito a um tipo de informação que existe independente da ação do homem. Hjørland (2003, p. 93) descreve que os pesquisadores de CI “não são *experts* na interpretação de informações sobre estrelas, mas são, na sua maior parte, especialistas na informação documentada por astrônomos”. O autor apresenta esse exemplo, no qual a informação é entendida em um sentido mais amplo do que normalmente está implícito na CI, para fundamentar a seguinte afirmação:

O foco especial da CI é o conhecimento documentado produzido pelos seres humanos em algum tipo de documento de uso potencial para outros seres humanos. A luz das estrelas não é informação para a comunidade da CI, mas informações astronômicas como as produzidas e utilizadas por astrônomos são. Esta distinção pode parecer sutil, mas é importante, a fim de construir uma base teórica sólida para a OC (HJØRLAND, 2003, p. 93, tradução nossa¹⁹).

Nesse sentido, Hjørland (2007 a) considera apropriado falar-se em signos ao discutir aspetos da OC na CI. A teoria semiótica é considerada um quadro teórico de referência mais apropriado para os estudos da informação do que a teoria de Shannon e Weaver, por exemplo. Para o autor teorias objetivistas e universalistas da informação podem ter um apelo muito mais forte do que pontos de vistas teóricos que consideram informação, significado e decisões dependentes do contexto. No entanto, uma consideração superficial da natureza da informação nos deixaria sem uma fundamentação teórica adequada. (HJØRLAND, 2007 a).

Na tradição semiótica pragmática, desenvolvida pelo filósofo e lógico estadunidense Charles Sanders Peirce (1839-1914), conceitos são entendidos como signos. De acordo com Thellefsen; Thellefsen (2004, p.180, tradução nossa²⁰) “Isso nos permite afirmar que o vasto conteúdo

¹⁹ The special focus of LIS is on documented knowledge produced by human beings in some kind of documents of potential use to other human beings. Light from the stars is not information for the LIS community, but astronomical information as produced and used by astronomers is. This distinction may seem subtle, but is important in order to construct a firm theoretical basis for KO.

²⁰ This allows us to maintain that the vast content of tacit knowledge within a knowledge domain expresses stable patterns of meaning, which exactly is the result of sign processes.

de conhecimento tácito de um domínio de conhecimento expressa padrões estáveis de significação, que é resultado de processos sgnicos”. Estes processos sgnicos (semiose) posteriormente tornam-se hábitos de interpretações.

De acordo com a nossa compreensão do conhecimento, a interpretação dos signos em um contexto social não depende apenas do indivíduo intérprete, mas está ancorada no contexto social. Não se pode interpretar os signos de acordo com caprichos subjetivos ou preferências. O domínio do conhecimento coloca restrições sobre as nossas interpretações. (THELLEFSEN; TTHELLEFSEN, 2004, p. 180, tradução nossa²¹)

Hjørland (2009, p.1522, tradução nossa²²) afirma que “conceitos são construídos dinamicamente e os significados, que classificam o mundo de acordo com interesses e teorias, são negociados coletivamente.” O autor trata a teoria do conceito sob ótica epistemológica e considera as visões pragmática e historicista como as mais produtivas para o desenvolvimento da OC. As abordagens pragmática e historicista na OC enquadram-se no paradigma social da CI.

Para Hjørland (2009, p. 1530, tradução nossa²³), “conceitos tem sido entendidos como significados socialmente negociados”. Essa negociação e, conseqüentemente, a formação dos conceitos, por se tratar de um processo social, ocorrem em contextos múltiplos e heterogêneos, que devem ser levados em conta, uma vez que deles dependem²⁴.

²¹ According to our understanding of knowledge, the interpretation of signs within a social context is not solely dependent on the individual interpreter but is anchored within the social context.

²² Concepts are dynamically constructed and collectively negotiated meanings that classify the world according to interests and theories.

²³ Concepts have been understood as socially negotiated meanings that should be identified by studying discourses rather than by studying individual users or a priori principles.

²⁴ Além dos trabalhos citados ao longo dessa seção destacamos Almeida (2011), Frohmann (1992), Mostafa (2012) e Siqueira (2008) como trabalhos que também contribuíram na formulação do enunciado 1.

4.1.2 Contextos múltiplos e heterogêneos e dilemas éticos

Vários estudos da CI, como os bibliométricos e de OC, são desenvolvidos no contexto de determinado domínio do conhecimento. Isso se deve, dentre outros fatores, à necessidade de minimizar as diferenças conceituais. No entanto, diferenças conceituais e concepções concorrentes podem ocorrer mesmo no interior de determinada área do conhecimento.

A análise de domínio, bem como a abordagem pragmática em CI, assume que diferentes abordagens (ou paradigmas) existem em todos os domínios do conhecimento. De acordo com Hjørland (2008b), esses paradigmas devem ser identificados no processo de OC. No entanto, o autor esclarece que podem existir abordagens ou paradigmas não representativos, que não são apropriados para a OC. Dessa forma, qualquer SOC acaba apresentando um tipo de inclinação, ou *bias*²⁵, para os pontos de vistas mais representativos da área.

Outro aspecto da inevitável inclinação nos SOC, nos parece dizer respeito à linguagem em si, “as linguagens seccionam o mundo de maneiras diferentes e não dispomos de nenhum meio sub-linguístico neutro de informar.” (KUHN, 1975, p. 437, apud GONZALEZ DE GOMEZ, 2001, p. 8).

Para a autora, Wittgenstein colocou uma espécie de sombra sobre as concepções universalistas da linguagem, uma vez que, na sua visão, não existe nenhuma linguagem capaz de sintetizar outras linguagens. Os paradigmas de Kuhn fazem o mesmo sobre os territórios discursivos da ciência.

A aproximação entre o pensamento de Wittgenstein e Kuhn é percebida também nos diversos trabalhos de Birger Hjørland que fazem parte do corpus dessa pesquisa. Nos quais o autor aponta a importância de estudos epistemológicos para fundamentar a OC. O conceito de paradigma em seus trabalhos é ampliado, se aproximando da ideia de jogos de linguagem que diz respeito à linguagem comum ou ordinária e ultrapassa a barreira do discurso científico. Hjørland (2003, p. 100, tradução nossa²⁶) afirma que:

²⁵ Em outro trabalho, o autor citado aborda especificamente a questão do *bias* na OC, no qual assim como Melaine Feinberg sugere, considera-se que “se não podemos eliminar o *bias*, então devemos nos esforçar para ser mais responsáveis sobre ele e determinar explicitamente a perspectiva representada no sistema de informação”. (Hjørland, 2008a).

²⁶ “Paradigm” may, in the philosophy of science, mean more or less implicit background assumptions concerning the object of research, concerning research methods, concerning the usefulness of research, and so forth. In an even broader meaning the concept may be

“Paradigma”, na filosofia da ciência, pode significar suposições mais ou menos implícitas sobre o objeto de pesquisa, os métodos de investigação, a utilidade da pesquisa, e assim por diante. Em um sentido ainda mais amplo o conceito pode ser usado com pontos de vista mais ou menos coerentes sobre uma prática coletiva, como o ensino, o serviço religioso, a administração da justiça, a política, e assim sucessivamente. O proscrito de Kuhn na segunda edição publicada em 1970 define, entre outras coisas, um paradigma como uma constelação de crenças, valores, técnicas, entre outros compartilhados por membros de determinada comunidade.

Consideramos que para a OC essa ampliação do conceito de paradigma é importante, pois ultrapassa a barreira do campo científico e acadêmico. Abrange demais aspectos contextuais envolvidos na produção e uso do conhecimento. Andrade (2012, p. 27) afirma que “na concepção de Armengaud (2008), o contexto é um conceito central e caracterizante para a pragmática”.

Para Kobashi e Fernandes (2009 apud ANDRADE, 2012, p. 28) a “abordagem de Armengaud indica que é mais fácil reconhecer a existência de diferentes contextos do que delimitá-los operacionalmente para fins da organização da informação”.

Essa dificuldade em se delimitar os diferentes contextos, junto com a aparente necessidade de minimizar as diferenças conceituais decorrentes de aspectos contextuais que podem dificultar a comunicação, faz com que na OI e OC a delimitação do contexto siga a divisão científica e/ou acadêmica vigente.

No entanto, se a maneira como o conhecimento é desenvolvido fosse analisada profundamente, delimitações de contextos mais apropriadas poderiam ser encontradas. Gonzalez de Gomez (2001, p. 9), com base no pensamento de Foucault, afirma que:

used with more or less coherent views concerning a collective practice, such as teaching, divine service, administration of justice, politics, and so forth. Kuhn’s 1970 postscript in the second edition defines among other things a paradigm as a constellation of beliefs, values, techniques and so on shared by members of a given community.

A máquina epistemológica e suas operações de demarcação tiveram uma esfera de ação bem maior que os espaços formais das academias e as instituições de pesquisa. Ela teria agido nos laboratórios, mas também nas escolas, nos confessionários e nas bibliotecas, nos museus e nos hospitais. A operação de demarcação, ampliada nos domínios das instituições reguladoras do cotidiano, instaura o regime de verdade dominante. A partir dessa concepção, Foucault destacou-se como cuidadoso experimentador de metodologias para lidar com a singularidade dos acontecimentos prático-discursivos, especificamente, aqueles que por seu caráter estigmatizado ou periférico melhor revelariam as estratégias modeladoras dos regimes de verdade.

Alvarenga (1998) considera que o pensamento de Michel Foucault, em *A Arqueologia do Saber*, poderia trazer contribuições, possibilitando visões mais abrangentes dos campos de conhecimento e ensejando novas possibilidades de análises de discursos oriundos de instâncias diversas, não somente das restritas à ciência normal. O que vai ao encontro da visão de Saldanha (2008, p. 23) ao afirmar que:

Há no “paradigma emergente” de Santos (1987) uma aproximação ao senso comum, ou seja, a todo o conjunto de outros conhecimentos não necessariamente sediados nos campi acadêmicos. Trata-se de uma filosofia da linguagem ordinária, fundamento das Investigações Filosóficas de Wittgenstein (1979).

A extrapolação do campo científico é justificada por Alvarenga (1998, p. 3), pois segundo a autora:

A construção do conhecimento deve ser vista como um processo eivado de interesses pessoais e subjetividade, devido aos imperativos que governam a conduta social implícita na construção não somente da ciência, mas dos saberes em geral.

E uma vez que:

O conhecimento científico repousaria sobre um suporte institucional, sendo reforçado e acompanhado por outros estratos e práticas sociais, tais como a política, pedagogia, o sistema de comunicação do conhecimento, incluindo os sistemas de editoração e bibliotecas. [...] a biblioteca seria vista como uma instituição para arranjar textos, estabelecendo-se como um componente da legitimação de uma ordem particular do discurso. (ALVARENGA, 1998, p.3).

Este arranjo de textos, que é um componente de legitimação, precisaria levar em conta a questão da dispersão.

A categoria *dispersão*, no âmbito da ciência da informação, vincular-se-ia à constatação de núcleos, de seleções, resultando em produtos, tais como os formados de autores considerados mais produtivos, ou os que se constituem de listas básicas de periódicos, as “*core lists*”. A *dispersão* é normalmente analisada a partir de um universo composto, segundo um princípio de seleção e os resultados nem sempre são complementados pelos elementos excluídos dessas seleções, fato que pode redundar em conclusões incompletas ou mesmo equivocadas.

Em Foucault, a *dispersão* tende a sugerir que não se deve pautar um estudo arqueológico justificado por opções somente calcadas em princípios de seleção para que não se construam grupos a partir somente de privilégios (autores mais produtivos, periódicos mais citados etc.), descartando-se os elementos não privilegiados. Haveria de se considerar também nas análises e discussões os segmentos excluídos das seleções. (ALVARENGA, 1998, p.4 e 5).

Princípios de seleção são comumente empregados na OC, Hjørland (2012) considera que a questão básica na classificação é vista como o fornecimento de critérios para decidir se algo deve ser classificado como X. Nesse trabalho o autor supracitado analisa os desafios enfrentados na classificação bibliográfica, tanto no nível teórico quanto no prático. No nível prático, as bibliotecas estariam cada vez

mais dispensando a classificação de livros. No plano teórico, muitos pesquisadores, gestores e usuários acreditam que a atividade de “classificação” não valeria mais a pena devido ao esforço, pois consideram que os motores de busca podem fornecer meta-dados de forma mais fácil.

Google e outros sistemas de RI semelhantes certamente são impressionantes, mas como eles classificam e priorizam as informações relevantes? Temos tendência a pensar em tais sistemas como ferramentas neutras e objetivas, mas não pode ser. Qualquer sistema está sempre inclinado [*biased*] de uma maneira ou de outra (ver, por exemplo, Fortunato et al., 2005; Gerhart, 2004; Introna e Nissenbaum, 2000). Os motores de busca podem ser calibrados de forma a proporcionar diferentes resultados ou rankings. A fim de fazer essa calibração (ou simplesmente para avaliar os sistemas), é preciso ter algum tipo de classificação do que deve ser encontrado. Até agora, no campo da CI, temos utilizado principalmente as avaliações de relevância baseadas na “relevância do usuário” (ver Hjørland, 2010). (HJØRLAND, 2012, p. 311, tradução nossa²⁷).

Em Hjørland (2010), por sua vez, encontramos maiores informações quanto à forma de seleção e a determinação da relevância.

Os motores de busca da Internet, como o Google baseiam-se em algoritmos. Um motor de busca identifica alguns documentos quando um usuário digita alguns termos de pesquisa. Na realidade, é claro, alguns programadores desenvolveram um algoritmo que determina qual deve ser a

²⁷ Google and other similar IR systems are certainly impressive, but how do they classify and prioritize the relevant information? We tend to think of such systems as neutral and objective tools, but they cannot be. Any system is always biased in some way or another (see, for example, Fortunato et al., 2005; Gerhart, 2004; Introna and Nissenbaum, 2000). Search engines may be calibrated in order to provide different findings or rankings. In order to make such a calibration (or simply to evaluate the systems), we need to have some kind of classification of what should be found. Thus far, in the field of LIS, we have mainly used assessments of relevance based on “user relevance” (see Hjørland, 2010).

prioridade dada aos documentos. É, portanto, não o algoritmo ou o sistema que determina a relevância, mas os programadores humanos (sendo intencional ou não). Neste contexto, é importante considerar a afirmação de Saracevic: "O ponto de toda avaliação de diferentes algoritmos é que eles produzem saídas diferentes para a mesma consulta e para o mesmo conjunto de documentos no sistema" (Saracevic, 2007, p. 1930).

A seleção do sistema (ou seja, do programador) não é, portanto, "Perfeita" ou "objetiva", mas é uma escolha feita entre muitas escolhas possíveis. Pode ser mais ou menos inteligente e mais ou menos adequada para diferentes fins. Portanto, é subjetiva (E é difícil entender por que tantas pessoas afirmam o contrário). Sua subjetividade é determinada pelas escolhas do programador, incluindo opções de pontos de acesso de assunto, de ponderação, e utilização de estruturas de ligação. Além disso, e talvez o mais importante: A escolha dos métodos de avaliação. (HJØRLAND, 2010, p. 218, tradução nossa²⁸)

O autor prossegue a sua argumentação e descreve a diferença entre os conceitos de “objetivo” e “subjetivo”. “Uma instrução ou representação é subjetiva se remete às opiniões, crenças e sentimentos de convicção deste ou daquele indivíduo. É objetivo se é independente

²⁸ Internet search engines such as Google are based on algorithms. A search engine identifies some documents when a user enters some search terms. In reality, of course, some programmers have designed an algorithm that determines what documents should be given priority. It is thus not the algorithm or the system that determines relevance, but the human programmers behind the engine (whether they intended so or not). In this connection, it is important to consider Saracevic's statement: "The whole point of the evaluation of different algorithms is that they produce different outputs for the same query and from the same set of documents in the system" (Saracevic, 2007, p. 1930). The system's (i.e., the programmer's) selection is thus not "perfect" or "objective" but is a choice made among many possible choices. It may be more or less clever and more or less suited to different purposes. Therefore, it is subjective (and it is difficult to understand why so many people claim the opposite). Its subjectivity is determined by the programmer's choices, including choices of subject access points, weighting, and utilization of link structures. Also, and perhaps most important: The choice of evaluation methods.

da opinião das pessoas, crenças e pontos de vista”. (HJØRLAND, 2010, p.218, tradução nossa²⁹).

Hjørland (2013b) considera que SOC não podem ser objetivos e devem, portanto, por implicação, conter algum tipo de subjetividade. Esta subjetividade deve, no entanto, ser derivada de opiniões coletivas nas comunidades discursivas, em vez de ser derivada de estudos de indivíduos ou a partir do estudo abstrato da mente.

Frohmann (2008b, p. 267), considera que a subjetividade está no centro da questão da ética da informação. Investigando o significado e o papel da subjetividade na ética da informação propõe respostas para algumas questões, “como a subjetividade é entendida? Pode ser entendida de forma que amplie a reflexão ética para incluir problemas que permanecem invisíveis quando a subjetividade é tida como certa e quando a forma como é criada permanece inquestionável?”

Para desenvolver sua argumentação o autor apresenta a contribuição de diversos autores (Froehlich, Hauptman, Foucault, Deleuze e Hacking). Citando Gilles Deleuze, o autor afirma que:

Num espírito foucaultiano, ele argumentou que a ética da informação precisa ser genealógica, porque é preciso olhar para as possibilidades morais e éticas embutidas nos regimes de informação³⁰ que habitamos.[...] a individualidade e subjetividades que são possíveis em qualquer momento histórico estão vinculadas com os sistemas sócio-técnicos que as tornam possíveis. A ética da Informação deve levar em conta as possibilidades da ação ética em zonas onde os sujeitos não são totalmente determinados, mas onde as forças de estabilização, codificação, territorialização, e dominação no trabalho de construir e configurar os modos de subjetivação possam sofrer negociação, oposição, resistência, e transmutação. Na medida em que a ética está preocupada com a subjetividade, o problema não está localizado em debates sobre como

²⁹ A statement or a representation is subjective if it refers to the opinions, beliefs, and feelings of conviction of this or that individual. It is objective if it is independent of people's opinions, beliefs, and views.

³⁰ Chamaremos ‘regime de informação’ a um sistema ou rede mais ou menos estável na qual a informação flui através de canais determináveis - de produtores específicos, via estruturas organizacionais específicas, a consumidores ou usuários específicos. (FROHMANN, 1995, p.5-6).

determinados sujeitos morais devem agir, mas na interação de poder entre as forças de dominação e as possibilidades de liberdade na formação de sujeitos. Precisamos de uma ética da informação que reconhece como os processos e tecnologias de informação estão envolvidos no desenvolvimento humano. (FROHMANN, 2008b, p. 273, tradução nossa³¹)

Para Frohmann (2008b, p. 275) “Um estudo abrangente da importância de Foucault para a ética da informação ainda precisa ser feito”. Assim, procura mostrar que a ética da liberdade de Foucault oferece uma crítica à ética da informação *self-centered*. O autor explica que,

A diferença entre a ética da liberdade de Foucault e a ética *self-centered* de Froehlich e Hauptman é que no último caso, as subjetividades são muito duráveis e muito estáveis. Apesar de serem componentes emergentes de *assemblages* morais, suas configurações são tomadas como dadas de uma vez por todas, sem a possibilidade de suas contingências ou o fracasso de suas *assemblages* jamais serem recebidas. (FROHMANN, 2008b, p.276, tradução nossa³²).

³¹ In a Foucauldian spirit, he argued that information ethics needs to be genealogical because “we need to look at the moral and ethical possibilities embedded in the regimes of information that we inhabit” (p. 299). [...] “the individuality and subjectivities that are made possible at any historical time are bound with the socio-technical systems that make them possible” (p. 299). Information ethics must take account of the possibilities of ethical action in zones where subjects are not fully given, but where the forces of stabilization, coding, territorialization, and domination at work to construct and configure modes of subjectivation can be negotiated, opposed, resisted, and transmuted. Insofar as ethics is concerned with subjectivity, the problem is not located in debates about how given moral subjects ought to act, but in the interplay of power between forces of domination and possibilities of freedom in the formation of subjects.¹⁵ We need an information ethics that acknowledges how information processes and technologies are implicated in making up people.

³² The difference between Foucault’s ethics of freedom and the self-centered ethics of Froehlich and Hauptman is that in the latter case, the subjectivities are *too* durable and *too* stable. Although they are emergent constituents of moral assemblages, their configurations are taken as given once and for all, without the possibility of their contingencies or the failure of their assemblages ever being entertained.

Assemblages, segundo Frohmann (2008b, p. 275, tradução nossa³³) se referem às,

Configurações em que muitos elementos heterogêneos (físicos, conceituais, discursivos, tecnológicas, institucionais) estão alinhados para produzir o todo. Os elementos específicos que compõem *assemblages* variaram de caso para caso [...].

O conceito de *assemblage*, especialmente elaborado por Latour, ajuda-nos a situar os documentos e a documentação em associações complexas de entidades amplamente heterogêneas. Também nos redireciona para a ética e a política dos estudos de documentação que são sensíveis às forças de associação ou reunião, que são tão importantes para a análise de Latour. Ele nos mostra como a documentação pode ser uma questão de ética e política importante de preocupação para a concepção de formas de vida coletiva. (FROHMANN, 2008a, p.179, tradução nossa³⁴).

Nesse sentido, completamos a explicação sobre como a ética da liberdade de Foucault pode trazer contribuições para a ética da informação,

Foucault, Deleuze e Hacking nos mostram uma forma de pensamento ético que ilumina um campo de problemas e questões em ética da informação veladas a partir da perspectiva de uma ética da informação *self-centered*. Seus trabalhos também nos confrontam com possibilidades de fuga, não só dos modos de dominar a subjetivação imanente aos sistemas de informação, mas de uma estreita

³³ Configurations in which many heterogeneous elements (physical, conceptual, discursive, technological, institutional) are aligned to produce wholes.

³⁴ The concept of assemblage, especially as elaborated by Latour, helps us situate documents and documentation in complex associations of widely heterogeneous entities. It also redirects us to the ethics and politics of documentation studies that are sensitive to the powers of association or gathering, which are so important to Latour's analysis. He shows us how documentation can be an important ethical and political matter of concern for designing ways of collective living.

faixa de problemas morais que vem junto com eles. Mas também sugere uma crítica: na medida em que uma ética da informação *self-centered* contribui para a estabilização dos modos de subjetivação e *assemblages* da qual emergem, é cúmplice da reificação das forças de dominação das quais a sua ética quer nos libertar. (FROHMANN, 2008b, p. 276, tradução nossa³⁵).

Sudin e Johannisson (2005) sugerem que um ponto fraco do neo-pragmatismo seria a sua insuficiência para analisar o conceito central de justificação em relação ao poder. E enfatizam a importância de levar em consideração o poder quando se estuda necessidade, uso e busca de informação.

Por exemplo, os textos de Richard Rorty dão muito pouca atenção a uma problematização de como as reivindicações de certos grupos de conhecimento têm precedência na sociedade comparada com a de outros ou por que certas reivindicações de conhecimento, são excluídas. [...] um interesse insuficiente no poder é mais ou menos evidente no trabalho de estudiosos da CI que têm vindo a utilizar uma epistemologia pragmática. (SUDIN; JOHANNISSON, 2005, p.40, tradução nossa³⁶).

No entanto, consideramos que as pesquisas em CI apresentadas nesse item, bem como no item 4.1.3 a seguir, não se enquadram nesse perfil. Pois essas apresentam interesse nas questões relacionadas ao poder e à influência que as estruturas científicas e sociais vigentes

³⁵ Foucault, Deleuze, and Hacking show us a way of ethical thinking that illuminates a field of problems and issues in information ethics veiled from the perspective of a self-centered information ethics. Their work also confronts us with possibilities of escape not only from dominating modes of subjectivation immanent to information systems but from a narrow range of moral problems that goes along with them. But it also suggests a criticism: That insofar as a self-centered information ethics contributes to the stabilization of modes of subjectivity and the assemblages from which they emerge, it is complicit in the reification of the forces of domination from which their ethics seeks to free us.

³⁶ For example, Richard Rorty's texts give very little attention to a problematization of how the knowledge claims of certain groups are given precedence in society compared to that of others or why certain knowledge claims are excluded.[...] an insufficient interest in power is more or less evident in the work of LIS scholars who have been using a pragmatist epistemology

exercem sobre a produção e uso do conhecimento. Nesse contexto destacamos o pensamento de Michel Foucault, que baseou grande parte das pesquisas usadas na formação desse item, pois suas teorias abordam a relação entre poder, conhecimento e controle social.

Sobre o pensamento de Foucault, Hjørland (1998a) se refere a este na OC ao afirmar que é importante um tipo de análise relacionada ao “construtivismo social”, uma espécie de escavação de camada após camada (o que Michel Foucault chamou de “a arqueologia do saber”). Para Hjørland (2003, p. 94) a visão do construtivismo social “está relacionada com a visão pragmática [...]. Muitas vezes, porém, o construtivismo social e pragmatismo se opõem ao tipo de realismo, como realismo científico.” Mas a própria questão do realismo na visão pragmática é tema polêmico.

De acordo com Rorty, realismo científico e pragmatismo são dois pontos de vista que não podem ser combinados; pragmatismo é visto como uma posição anti-realista. De acordo com outros filósofos (Eg, Dewey, 1929; Ellis, 1990), o realismo deve ser a base do pragmatismo. Isso corresponde à minha própria visão, o que pode, assim, ser denominado realismo pragmático. Conceitos e campos científicos tendem a representar partes da realidade de uma forma que é funcional para a atividade humana. Na medida em que o construtivismo social, opõe-se ao realismo, não concordo com essa visão. Deve-se evitar a armadilha do reducionismo sociológico. No entanto, tenho encontrado muitas pesquisas feitas sob a bandeira do construtivismo social profundamente relevantes para a compreensão da estrutura de diversas áreas do conhecimento. (HJØRLAND, 2003, p. 94, tradução nossa³⁷).

³⁷ Pragmatism are two views that cannot be combined; pragmatism is seen as an antirealist position. According to other philosophers (e.g., Dewey, 1929; Ellis, 1990), realism must indeed be based on pragmatism. This corresponds to my own view, which may thus be termed *pragmatic realism*. Scientific concepts and fields tend to represent parts of reality in a way which is functional for human activity. To the degree that social constructivism is opposed to realism, I disagree with that view. One should avoid the pitfall of sociological reductionism. However, I have found much research done under the banner of social constructivism deeply relevant for the understanding of the structure of many knowledge fields.

Sudin e Johannisson (2005) afirmam que uma das críticas centrais ao neo-pragmatismo é a acusação de relativismo, apresentada a partir de uma posição realista. Mas essa acusação não é considerada válida, uma vez que a discussão acaba em uma espécie de beco sem saída, pois discórdias entre realismo e anti-realismo não poderiam ser resolvidas com a ajuda de critérios desenvolvidos no seio da comunidade científica. Os autores afirmam que “a questão ontológica se existe ou não realidade raramente é frutífera na CI ou outras ciências sociais”. (SUDIN; JOHANNISSON, 2005, p. 40).

Se uma realidade, que é exterior à mente humana, existe ou não, não podemos afirmar. O fato é que, na visão pragmática, a interpretação humana dos fenômenos é levada em conta para se entender a construção do conhecimento. Assim como as pesquisas no construtivismo social, linha na qual o pensamento de Foucault pode ser enquadrado. O pensamento de Foucault bem como a linha de estudos do construtivismo social é considerada útil para a OC apesar de o construtivismo social ser anti-realista, e o anti-realismo não ser bem aceito na visão pragmática. Hjørland (2004) argumenta a favor do realismo na CI, e inclui a abordagem pragmática na área como integrante da linha realista.

Frohmann (2008b, p. 275, tradução nossa³⁸) citando Latour (2005) nos ajuda a compreender como o construtivismo social pode ser compatível com o realismo.

O absurdo de supor que mostrar que algo é construído é para diminuir a sua realidade ou para mostrar que é uma farsa é execrado por Latour em seu livro de 2005, no qual ele diz: “em todos os domínios, dizer que algo é construído tem sido sempre associado à valorização de sua robustez, qualidade, estilo, durabilidade, valor, etc. Tanto

³⁸ The absurdity of supposing that to show something is constructed is to diminish its reality or to show it is a fake is excoriated by Latour in his 2005 book, where he says: “In all domains, to say that something is constructed has always been associated with appreciation of its robustness, quality, style, durability, worth, etc. So much so that no one would bother to say that a skyscraper, a nuclear plant, a sculpture, or an automobile is ‘constructed.’ This is too obvious to be pointed out” (p. 89); “Facts were facts—meaning exact—*because* they were fabricated—meaning that they emerged out of artificial situations. Every scientist we studied was proud of this connection between the quality of its construction and the quality of its data. This strong connection was actually one’s main claim to fame. While the epistemologists might have forgotten this, etymology was there to remind everybody” (p. 90); “When we say that a fact is constructed, we simply mean that we account for the solid objective reality by mobilizing various entities whose assemblage could fail”.

que ninguém se deu ao trabalho de dizer que um arranha-céu, uma usina nuclear, uma escultura, ou um automóvel são construídos. Isso é obvio demais para ser apontado”. [...] “quando dizemos que um fato é construído, significa simplesmente que representa a realidade objetiva e sólida mediante a mobilização de várias entidades cuja *assemblage* pode falhar.”

O fato do desenvolvimento do conhecimento depender da ação humana, e cada indivíduo ser influenciado por alguns pressupostos, concepções e influências teóricas, que podem trabalhar por meio da linguagem e outros fenômenos culturais, ou seja, é uma influência coletiva, possui uma relação de causalidade com a temática abordada na seção a seguir³⁹.

4.1.3 A qualidade da informação não é constante

A questão da qualidade da informação foi estudada por Assis e Moura (2011). As autoras buscaram responder se,

As abordagens convencionais do conceito de qualidade da informação atendem a um contexto caracterizado pela interatividade, colaboração e trocas simbólicas constantes mediadas pela arquitetura de redes sociais no qual o usuário passa a atuar como sujeito que não apenas utiliza, mas também produz, remodela e qualifica os conteúdos informacionais através da linguagem (ASSIS; MOURA, 2011, p. 97).

E observaram na literatura que as pesquisas não consideram a qualidade da informação como um processo, o que deveria ser reconsiderado uma vez que “nem informação nem qualidade são fenômenos constantes por que se modificam o tempo todo” (BRIER, 2006 apud ASSIS; MOURA, 2011, p. 16). Assim, considera-se que os modelos de qualidade da informação compõem um recorte arbitrário e operacionalizável de um universo multidimensional e complexo de

³⁹ Além dos trabalhos citados ao longo dessa seção destacamos Ekbja e Evans (2009), Gonzalez de Gomez (2005) e Mostafa (2010) como trabalhos que também contribuíram na formulação do enunciado 2.

significações. São criados para contextos e propósitos específicos e desse modo, propensos a sofrerem desatualizações (ASSIS; MOURA, 2011).

Hjørland(2000) aborda a questão da avaliação do conhecimento na visão pragmática, na qual os contextos e o histórico são considerados, bem como a importância dos documentos para essa contextualização.

Cada indivíduo é influenciado por alguns pressupostos, concepções e influências teóricas. Isto é verdade tanto para cientistas da informação, usuários, e todas as outras pessoas (em uma compreensão mais ampla até mesmo para os animais). Tais influências podem trabalhar através da linguagem e outros fenômenos culturais, e, portanto, a influência é mais ou menos coletiva. Cada indivíduo tem, assim, uma certa maquiagem teórica, que é formada por alguns contextos específicos e de alguns pressupostos. Esta composição influencia as maneiras pelas quais todos os textos são escritos, bem como a forma como eles são lidos. Também influencia a forma como os sistemas e serviços de informação são concebidos e geridos. A produção e avaliação do conhecimento não podem ser feitas apenas por princípios empiristas ou racionalistas, mas por uma combinação, com o conhecimento histórico da origem das teorias e considerando os objetivos e valores humanos. Conhecimento se torna mais contextualizado quando o analisamos sob a perspectiva dos documentos e de seu conteúdo. Um documento tem uma história, um ou mais autores ou produtores, uma conexão com outros documentos, e assim por diante. (HJØRLAND, 2000, p. 34 e 35, tradução nossa⁴⁰).⁴¹

⁴⁰ Each individual is influenced by some presumptions, conceptions, and theoretical influences. This is true both for information scientists, the users, and all other people (in a broader understanding even for animals). Such influences can work through language and other cultural phenomena, and thus influence in a more or less collective way. Each individual has thus a certain theoretical makeup, which is shaped by some specific contexts and from some assumptions. This makeup influences the ways in which all texts are written as well as how they are read. It also influences the way information systems and services are designed and managed. The production and evaluation of knowledge cannot be done by empiricist or rationalist principles alone, but by a combination, in

Hjørland (2011a) argumenta que, para avaliar uma fonte de informação, por exemplo, um verbete da Wikipédia, é necessário relacionar o conteúdo dessa fonte com a interpretação do estado do conhecimento na frente de pesquisa, que normalmente é mais desenvolvido. Para o mesmo autor, as fontes de pesquisa devem ser avaliadas em relação à forma como abordam as controvérsias e diferentes pontos de vista.

O caso de estudo no qual o artigo foi baseado se trata da comparação da qualidade do conteúdo sobre o rastreio do câncer de mama da Wikipédia, da Wikipédia em dinamarquês, da Enciclopédia Britânica e da Enciclopédia Nacional da Dinamarca. O rastreio do câncer de mama é um tema polêmico na pesquisa médica. O que segundo o autor:

Levanta a questão de saber se a metodologia sugerida só é relevante quando não há "conhecimento estabelecido". Uma razão para se concentrar em controvérsias é que os "fatos estabelecidos" (por exemplo, o ponto de fusão de chumbo) pode não ser adequado para diferenciar a qualidade das fontes de informação (por exemplo, quase todas as enciclopédias de hoje relatam corretamente o ponto de fusão do chumbo, variando somente sobre o número de casas decimais fornecidas). Eu também acredito que "o conhecimento estabelecido" é um termo problemático devido ao princípio de falibilidade. Portanto, eu prefiro falar sobre os graus de consenso e da natureza mais ou menos dinâmica de frentes de pesquisa. Eu acredito que o consenso é relativamente raro. Como Broadfield (1946) escreveu: "O consenso é mais provável que apareça entre os ignorantes, de quem é característico ser unânime quanto à verdade daquilo que é falso. Em questões intelectuais o

addition to historical knowledge of the origin of the theories and in considering human goals and values. Knowledge becomes much more contextualised, which brings the documents and their content into focus. A document has a history, one or more authors or producers, a connection to other documents, and so on.

⁴¹ "O ponto de vista do conhecimento como fatos isolados ou idéias está relacionado com o empirismo e o racionalismo, enquanto epistemologia pragmática olha para o conhecimento como um conjunto de teorias que cumprem algum propósito" (HJØRLAND, 2000, p. 34).

acordo é raro, especialmente em questões da vida "(p.69-70). Portanto, considero que a metodologia apresentada neste artigo é amplamente aplicável. (HJØRLAND, 2011a, p. 1897, tradução nossa⁴²).

O autor prossegue explicando que o método sugerido para avaliação de fontes de informação em si deve ser objeto de discussão e investigação na CI. Mas muitos cientistas da informação estão relutantes em se envolver em questões de descobertas científicas em outros domínios,

Evitar o envolvimento nestas questões, no entanto, pode resultar em resultados problemáticos em CI. Não é possível, por exemplo, avaliar uma fonte de informação, sem considerar os resultados específicos de um assunto na pesquisa contemporânea. (HJØRLAND, 2011a, p. 1897).

O envolvimento dos profissionais da CI que o autor propõe, em nossa opinião, seria não ignorar ou considerar como mais duvidoso um conteúdo que gere controvérsias. Pois a postura dos profissionais da informação perante essas questões se reflete no seu trabalho, que, por sua vez, influencia no desenvolvimento do conhecimento como um todo, uma vez que unidades de informação ocupam um papel importante na estrutura acadêmica e científica na qual esse desenvolvimento normalmente ocorre.

Ao se entender que a qualidade da informação não é constante, na nossa opinião, o papel do profissional da informação é ajudar a fornecer

⁴² Raises the question of whether the suggested methodology is only relevant when there is no “established knowledge.” One reason to focus on controversies is that “established facts” (e.g., the melting point of lead) may not be well suited to differentiate the quality of information sources (e.g., almost all encyclopedias today report the correct melting point of lead—only varying on the number of decimals they provide). I also believe that “established knowledge” is a problematic term because of the principle of fallibility. Therefore, I prefer to speak about degrees of consensus and the more or less dynamic nature of research fronts. I believe that consensus is relatively rare. As Broadfield (1946) wrote: “Consensus is most likely to appear among the unenlightened, of whom it is characteristic to be unanimous on the truth of what is false. In intellectual matters agreement is rare, especially in live issues” (pp. 69–70). Therefore, I consider the methodology presented in this article to be widely applicable.

os meios para que a sua avaliação se desenvolva da forma mais justa, equilibrada e democrática⁴³.

4.1.4 A prática é importante para a construção do significado

Frohmann (2004), para argumentar como a construção do fenômeno da informação pode ser estudado, apresenta contribuições de Wittgenstein. Pois considera que a investigação de Wittgenstein sobre “o que é o significado?” é importante para estudos da informação. Descreve alguns exemplos simples, extraídos do livro *Investigações Filosóficas*, que considera contribuir para deslocar nossa atenção das imagens mentalistas de significado para as práticas com a linguagem.

Com base em Wittgenstein, Hjørland (1998 a, p. 21, tradução nossa⁴⁴) afirma que:

Significados são produzidos por nossas práticas sociais. Uma consequência da prática social é o desenvolvimento da comunicação, do comportamento verbal e não-verbal e de conceitos. Os significados são produzidos em primeiro lugar “fora da mente” e são então, através da linguagem, transferidos para as mentes individuais. A partir dessa perspectiva, a questão central da semântica não está mais relacionada aos objetos ou mentes individuais, mas às culturas, às subculturas, à divisão social do trabalho, ao discurso das comunidades, às disciplinas científicas, e assim por diante.

Assim, o autor considera que as teorias mais adequadas sobre semântica seriam a sociocognitiva e sociolinguística e não apenas cognitiva e linguística. E no contexto mais amplo, da teoria sociocognitiva e sociolinguística, o autor aponta para as tradições

⁴³ Além dos trabalhos citados ao longo dessa seção destacamos Brothman (2010) e Mostafa (2011) como trabalhos que também contribuíram na formulação do enunciado 3.

⁴⁴ meaning are produced by our social practices. A consequence of social practice is the development of communication, of verbal and non-verbal behaviour, and of concepts. Mean-ings are first produced "outside the head" and are then, through language, transferred into the individual minds. From this perspective, the central question about seman-tics is not foremost related to individual objects or to in-dividual minds, but to cultures, to subcultures, to the so-cial division of labour, to discourse communities, to sci-entific disciplines, and so on.

pragmáticas⁴⁵ em semântica que precederam a teoria dos jogos de linguagem. Mais especificamente o pensamento de John Dewey e de Peirce.

Para John Dewey línguas são apenas um meio de comunicação de significado. Comunicação não-verbal, arte e objetos são todos expressivos; eles carregam significado, e podem ser considerados como uma espécie de linguagem. Cada arte tem o seu próprio meio e este meio é especialmente equipado para um tipo de comunicação. As necessidades da vida diária atribuíram mais importância ao discurso como forma de comunicação, em detrimento de outras. Diferentes culturas e necessidades humanas desenvolvem suportes especiais para comunicar significados. Para mim, esse ponto de vista parece intimamente relacionado com a teoria de Wittgenstein de “jogos de linguagem” (HJØRLAND, 1998 a, p.22, tradução nossa⁴⁶).

Sobre o pensamento de Peirce, Hjørland (1998 a, p. 21) afirma que ele “descobriu que a teoria pragmática do significado é “futurista” interpretando o significado do ponto de vista de como a determinação do sentido pode contribuir para o cumprimento das metas”.

Em Wittgenstein e em Peirce considera-se que o significado vem da prática social.

O referente é uma entidade que supõe a *práxis* social, como discutida por Blikstein (1995), à luz da análise do caso de Kaspar Hauser, e essa *práxis* é formada na relação que o sujeito estabelece com o mundo e com os outros, na realidade. Desse modo, o significado de algo não está relacionado apenas com o sistema de signos linguísticos

⁴⁵ Esse trabalho foi publicado em 1998, só em 2007 o autor, baseado nas pesquisas epistemológicas de Peregrin, veio a considerar a pragmática um paradigma em semântica.

⁴⁶ For John Dewey languages are only one medium of the communication of meaning. Non-verbal communication, art, and objects are all expressive; they carry meaning, and can be regarded as a kind of language. Each art has its own medium and that medium is especially fitted for one kind of communication. The needs of daily life have given superior practical importance to one mode of communication, that of speech. Different human cultures and needs develop special media to communicate meanings. To me, this view seems closely related to Wittgenstein's theory of "language games",

comunitariamente ligados, mas depende também das relações que os homens estabelecem com as coisas significadas. Não conhecemos as coisas na realidade efetivamente sem um contato com o referente, a realidade fenomênica, assim, os significados produzidos pelo processo de indexação e uso da LD, dependem de uma aproximação colateral com o referente. (SOUZA; ALMEIDA, 2012, p. 28).

O pragmatismo e o realismo de Charles Sanders Peirce (que integram a sua teoria semiótica), por levar em conta a condição de referencialidade é considerado por Almeida (2012c) uma opção teórica na OC para substituir o extremo relativismo. Almeida (2012a, p. 53) apresenta uma revisão da máxima do pragmatismo na qual Peirce (2000, p. 195) afirma que:

A fim de determinar o significado de uma concepção intelectual, dever-se-ia considerar quais consequências práticas poderiam concebivelmente resultar, necessariamente, da verdade dessa concepção; e a soma destas consequências constituirá todo o significado da concepção.

Mas Almeida (2012c, p. 207) também esclarece que:

A ação por si só não faz um conceito adequado, uma vez que o seu propósito depende da concepção criada antecipadamente e que dá sentido e referência para a ação, e não de outro modo. Como método, o pragmatismo visa "averiguar o verdadeiro significado de qualquer conceito, doutrina, proposição, palavra, ou outro sinal "(ALMEIDA, 2012c, p. 207).

Assim, certos tipos de concepções intelectuais não possuem significação pragmática. Almeida (2012a) apresenta o exemplo do conceito teológico de transubstanciação – a transformação do vinho em sangue e da hóstia em carne – que, ao aplicar-se o teste pragmático mostra-se não ter consequências práticas concebíveis. “Um conceito científico ou símbolo, por outro lado, respeitando as indicações do

objeto no processo de produção do interpretante, admite um exame pragmático.” (ALMEIDA, 2012a, p. 53).

Buscando aproximações do pragmatismo de Peirce com a OC Almeida (2012a, p. 53) afirma que:

A concepção de Peirce é um avanço em relação a abordagem de Dahlberg (1978ab), pois para chegar ao significado de um conceito não basta relacionar todos os predicados que aludem às principais características do objeto. Além disso, uma parte importante da definição é, como reconhece Dahlberg (1978a, p.14), considerar que um conjunto de características determina o conceito, contudo, pragmaticamente, ainda não diferenciaríamos intelectualmente as características dos objetos até que procedêssemos o teste pragmático, isto é, a busca por evidências práticas sensíveis e conceptualizadas mentalmente que confirmem os conteúdos dos juízos e dos predicados. Somente desse modo conheceríamos conceitos científicos pragmáticos e não simplesmente *legi-signos*, convenções sem referências necessárias aos objetos.

Os conceitos de um domínio de conhecimento, segundo Almeida (2010, p.18):

São considerados *signos* do tipo simbólico na perspectiva semiótica, pois são regidos pela convenção que estabelece que um conjunto de ideias se associa a um objeto. Nesse sentido, um símbolo supõe uma relação interpretativa com os objetos, a fim de gerar hábitos. Thellefsen (2002, p. 79) argumenta que a formação de hábitos nasce porque as pessoas, em um dado domínio do conhecimento, partilham metas comuns. A evolução dos conceitos em um contexto cultural para Thellefsen supõe a criação e a quebra de hábitos.

O conceito é produto da convenção e da formação de hábitos, mais do que resultado de simples qualidades, mesmo que estas caracterizem e individualizem o conceito em uma rede de

conceitos. Deste modo, o conceito é um símbolo, resultante da união entre signo e objeto. Os conceitos dependem da formação de hábitos, os quais estabelecem as regras de associação e de aproximação do *representamen*⁴⁷ com o objeto (ALMEIDA 2012 a, p. 52).

O papel das regras na teoria de Peirce, que dependem do hábito e aproximam o signo do objeto, é semelhante na teoria dos jogos de linguagem de Wittgenstein. Novelino (1998, p. 142) afirma que “o uso da linguagem, no sentido dado por Wittgenstein, implica no domínio e uso de regras, pois a aplicação correta de um termo significa que se age de acordo com as regras estabelecidas pelo contexto de sua aplicação.” As expressões linguísticas teriam significado apenas nos diferentes jogos de linguagem que são formações complexas de linguagem e ação.

As regras do jogo dizem respeito não somente aos signos linguísticos, mas, também, aos participantes da linguagem, aos objetos, às ações, ao contexto. Portanto, considerar o significado das expressões nos jogos de linguagem é considerá-lo do ponto de vista do modo de uso, isto é, como parte natural das ações humanas. O jogo de linguagem liga convencionalmente a expressão ao contexto no qual ela é proferida, fixando, assim, o modo específico de sua aplicação. Um significado de uma expressão pode ser assim considerado como sendo seu modo de uso, determinado pelas regras de um jogo de linguagem. (NOVELINO, 1998, p. 142).

Tanto no hábito de Peirce quanto no jogo de linguagem de Wittgenstein se fala em regras que definem a significação das expressões, ou seja, aproximam o signo do objeto representado por este. Gonzalez de Gomez (1996) considera que Redón Rojas (1996) mostra um dos temas relevantes para as pesquisas em biblioteconomia, as regras de estruturação para que os símbolos nos levem ao mundo da informação. Pois sem estas regras os documentos, de acordo com a autora, não nos dizem nada.

As regras não seriam representações mentais, mas sim práticas sociais que organizam a experiência prática e discursiva gerando regularidades que se constituem sobre o contrato local dos participantes

⁴⁷ Sinônimo de signo.

de uma relação intersubjetiva mediada pela linguagem (GONZALEZ DE GOMEZ, 1996).

Assim, a autora destaca a potencialidade da teoria da ação comunicativa de Habermas, na qual se diferencia o ato de enunciar (ilocucionário) do conteúdo enunciado (locucionário), permitindo combinações do dizer com o dito no ato de falar. Gonzalez de Gomez (2000, p.18) esclarece que a teoria da ação comunicativa é uma das versões contemporâneas das teorias da ação social. E “a ação social se assenta na força da relação de interlocução, que se estabelece ao mesmo tempo que a definição do vínculo social, conforme uma forma específica de comunicação e de transferência de informação.”

Gracioso (2008) também destaca a potencialidade dessa teoria, bem como sua ligação com o pensamento de Wittgenstein. Para a autora na abordagem pragmática na filosofia da linguagem existiria a ideia em comum entre a filosofia de Habermas e Wittgenstein de que a linguagem é uma prática. “Por isso, o significado da linguagem só poderia ser investigado a partir de seu uso objetivo, e não somente do abstrato.” (GRACIOSO, 2008, p. 15).

Gracioso (2008) esclarece que as abordagens filosóficas desses dois pensadores se contrapõem em muitos aspectos mas em seus planos pragmáticos de concepção da significação se complementam. Esta complementaridade é identificada e explicada de acordo com os propósitos da pesquisa da autora.

L.Wittgenstein é o filósofo que, em particular, a partir de sua obra *Investigações filosóficas*, ofereceu-nos recursos teóricos que nos explicaram as situações cotidianas simples sob as quais dependem o processo significativo. O autor nos propiciou teoricamente a fluidez característica dos processos de significação da linguagem atrelada ao seu uso. Obtivemos assim o desprendimento necessário para pensarmos a relação da linguagem e seu uso nos processos de significação frente a polifonia que caracteriza a *Web*. Não buscamos em Wittgenstein um mapa pontuando processos da significação e sim, reconhecer seus labirintos. [...]

Wittgenstein, a nosso ver, não pode ser pensado do ponto de vista de sua aplicação, mesmo porque, segundo o próprio filósofo, sua obra não apresenta nenhuma teoria ou metodologia. Ele

particularmente nos permitiu repensar os pré-conceitos que temos sobre a significação da linguagem. (GRACIOSO, 2008, p. 160 e 161).

Enquanto Habermas ofereceria um aporte teórico que reconhece e articula o processo de significação da linguagem em uso mais operacionalizável. Segundo a autora “foi a Pragmática formal de Habermas que nos indicou, dentro de uma consistente teoria comunicativa, as peças e as jogadas que precisariam ser articuladas até mesmo antes de se iniciar o uso da linguagem”. (GRACIOSO, 2008, p.161).

De acordo com Gracioso (2008, p. 155 e 156) para Habermas:

Haveria uma normativa no mundo da vida linguística estruturado (proposta inversa ao movimento dos jogos de linguagem de Wittgenstein). Essa normativa foi pensada por Habermas com o intuito de construir uma teoria sociológica da ação (via comunicação, uso da linguagem). A linguagem teria uma potência voltada a promover o entendimento mútuo na sociedade. A busca de entendimento entre sujeitos seria o combustível do desenvolvimento social e o uso da linguagem, a comunicação, o seu motor. A verdade, ou ainda, a validade das informações são estabelecidos argumentativamente e não apenas empiricamente.

No meio científico e acadêmico, principalmente nas chamadas ciências duras, esse estabelecimento é muitas vezes considerado apenas empírico. Mas cada vez mais a visão de Habermas, que em nossa opinião é compatível com o construtivismo social e o pensamento de Foucault abordado no item 4.1.2, tem se mostrado correta. No item a seguir, 4.1.5, prosseguimos com a discussão sobre esse tema⁴⁸.

⁴⁸ Além dos trabalhos citados ao longo dessa seção destacamos Almeida (2005; 2009), Frohmann (1992), Gonzalez de Gomez (2005), Redigolo e Almeida (2012) e Hjørland e Albrechtsen (1995) como trabalhos que também contribuíram na formulação do enunciado 4.

4.1.5 A interação é importante para a validação

Gracioso (2008) dando atenção às ações de busca de informação em ambientes que possibilitam a interação, a partir do reconhecimento do ator dessas ações de informação como agente de interação e interlocução, ator social e comunicativo, afirma que:

Os papéis desses intersujeitos, no âmbito das conexões transversais da rede, extrapolam as atividades operacionais de recuperação da informação. Esses adquirem um tipo de autonomia para busca e interação informacional que antes não lhes era oferecida na interface com bases referenciais sistematizadas. No cosmos semiótico das redes digitais, o agente da busca se depara com a necessidade de critérios e de subterfugios relacionados à confiabilidade e à segurança informacional. As ações de busca e de recuperação da informação não acontecem exclusivamente em contextos institucionalizados e organizados segundo metodologias sistêmicas formais, por isso, questionamos-nos como os intersujeitos das redes aceitam e validam propostas informacionais. (GRACIOSO, 2008, p.13).

A pesquisa reflete a preocupação com a validação no processo de busca por informação na *Web*, em que a qualidade da informação não é previamente avaliada, mas uma vez que, conforme já apresentamos nos itens 4.1.2 e 4.1.3, a forma de entender a qualidade da informação deveria ser repensada mesmo em ambientes institucionalizados (onde se desenvolvem os tesouros), consideramos pertinente apresentar os apontamentos teóricos desenvolvidos pela autora. Em sua pesquisa busca na teoria da ação comunicativa de Habermas, bem como em pesquisas baseadas nessa teoria, a resposta sobre como os sujeitos aceitam e validam a informação. Uma dessas pesquisas analisadas pela autora é a de Gerald Benôit que desenvolveu testes em sistemas informacionais e conclui que a aproximação dos critérios de significação construídos entre “humanos” é difícil de ser apropriada (*a priori*) completamente.

Guasque (2008), ao buscar identificar a relação entre o pensamento reflexivo proposto por John Dewey e as competências empregadas na busca e no uso da informação, verificou que:

- Ao se depararem com conhecimentos novos e pontos de vista diferenciados, pesquisadores em formação articulam e aplicam os critérios de autoridade e coerência da abordagem para avaliar a informação e suas fontes, analisando a estrutura e lógica dos argumentos ou métodos. [...]
- Pesquisadores em formação validam a sua compreensão e interpretação por meio de conversas com outros colegas, especialmente os do grupo de estudo e os professores. (GUASQUE, 2008, p. 202).

Estes resultados empíricos, a nosso ver, reforçam os argumentos teóricos de Gracioso (2008) que se baseiam na teoria da ação comunicativa de Habermas.

Com a existência de um processo interativo, de reestruturação e de construção de conteúdos, Gracioso (2008, p. 152) afirma que, “pode-se dizer que critérios de validação foram atingidos e, portanto, foi possível estabelecer um entendimento mútuo sobre as ações proferidas (o que não necessariamente pode resultar em acordos)”.

Sobre a questão de acordos Guedes (2010), em trabalho que busca insumos teóricos nos pressupostos do pensamento dialógico de Mikhail Bakhtin⁴⁹ (1895-1975) para elucidar a natureza das ações interdiscursivas recorrentes na indexação social, afirma que:

O diálogo deve ser visto como uma das manifestações do dialogismo. Fiorin (2006b, p. 24) explica que o vocábulo “diálogo”, portador do significado de “solução de conflito”, “entendimento”, “busca de acordo”, entre outros, “pode levar a pensar que Bakhtin é o filósofo da grande conciliação entre homens. Ao contrário, as relações dialógicas podem ser contratuais ou polêmicas, de divergências ou de convergências, de aceitação ou recusa [...]”. A ligação entre os

⁴⁹ “o principal tema presente nos textos do Círculo de Bakhtin era o estudo da linguagem no processo de interação social, o que mais tarde seria difundido pelo Círculo pelo termo dialogismo.” (GUEDES, 2010, p. 51).

discursos, isto é, o movimento interdiscursivo pregado pelo dialogismo, é sempre assimétrico, heterogêneo e conflituoso. (GUEDES, 2010, p.61).

Sendo assim, pode-se concluir que por meio da interação, pontos de vista são contrapostos e podem ou não ser validados, aceitos como verdadeiros. Na abordagem pragmática na OC considera-se que se existirem pontos de vistas diferentes, esses devem ser apresentados aos usuários. Na seção a seguir descrevemos como esses pontos de vista devem ser identificados para então serem apresentados aos usuários⁵⁰.

4.1.6 A bibliografia como fonte para identificar pontos de vista distintos

A abordagem pragmática em CI, conforme descrevemos ao longo dessa pesquisa, bem como a análise de domínio, desenvolvida por Birger Hjørland e Hanne Albrechtsen, assumem que diferentes abordagens (ou paradigmas) existem em todos os domínios do conhecimento. Assim, esses devem ser ativamente pesquisados na OC, pois “tais pontos de vista diferentes podem ser explícitos ou implícitos e, se forem implícitos, podem ser descobertos pela análise teórica e filosófica”. (HJØRLAND, 2007b, p.374, tradução nossa⁵¹).

Um item central para essa análise seria a bibliografia da área. Hjørland(2007b) retoma o princípio da garantia literária, cunhado em 1911 pelo britânico Edward Wyndham Hulme, o qual diz respeito a identificação e validação dos elementos dos SOC por meio da análise da bibliografia.

Qualquer bibliografia de um certo tamanho tem de enfrentar maneiras conflitantes de definição de conceitos e determinação de relações semânticas. Garantia literária não significa identificar apenas um texto do qual as relações semânticas podem ser inferidas. A tarefa é negociar entre diferentes critérios por a frente em diferentes textos, e selecionar a que tem o maior grau de autoridade

⁵⁰ Além dos trabalhos citados ao longo dessa seção destacamos Abrahamson e Rubin (2012), Gracioso (2010 a) e Ng (2005) como trabalhos que também contribuíram na formulação do enunciado 5.

⁵¹ Different views may be explicit or implicit. If they are implicit, they can be uncovered by theoretical and philosophical analysis.

cognitiva ou é considerado o melhor em relação ao objetivo do SOC. (HJØRLAND, 2007b, p. 388, tradução nossa⁵²)

Assim como este princípio de garantia literária, existem outras “garantias” empregadas na OC, para a validação dos elementos dos SOC, sobre as quais Gracioso (2008) afirma que:

Essas garantias agregavam valor aos processos sistêmicos e seletivos de busca de informação, aumentando as chances de sua relevância e pertinência para o usuário do sistema. [...] Contudo, no momento em que elementos culturais passam a ser garantias para representação do conhecimento, **implicações éticas** também precisaram ser consideradas. [...] questionamos se seria possível aplicar, aos instrumentos de representação temática, *garantias de uso* e *garantias culturais* que dessem conta da multiplicidade e da heterogeneidade das condições da informação virtualizadas a partir de tecnologias (GRACIOSO, 2008, p. 52 e 57, grifo nosso).

A garantia literária, provavelmente, também não é capaz de dar conta dessa multiplicidade e heterogeneidade. Contudo, quando pensamos na garantia literária, podemos dizer que talvez estas implicações estariam mais relacionadas com o desenvolvimento do conhecimento científico em si, e a consequente produção da bibliografia com base neste, do que com os instrumentos de representação temática. No item 4.1.9 voltamos a abordar esse tema.

Para Saldanha (2011) a afirmação do homem em suas relações sociais como horizonte das investigações – o homem e não a informação como objeto da Ciência da Informação (CAPURRO, 1992) – permite a identificação da hipótese inicial da relação entre o pragmatismo e a discursividade do humanismo no campo informacional.

⁵² Any bibliography of a certain size must confront conflicting ways of defining concepts and determining semantic relations. Literary warrant does not mean identifying only one text from which semantic relations may be inferred. The task is to negotiate between different claims put forward in different texts and to select the one that has the highest degree of cognitive authority or is considered best in relation to the goal of the KOS.

A unidade gramatical dos estudos pragmatistas da informação gira em torno da análise da forma como determinadas comunidades constroem, a partir das linguagens que comungam, suas apropriações sobre a informação. A comunidade discursiva específica, reconhecida em sua cultura, é o objeto de estudo; para penetrá-lo, é necessário, preliminarmente, conhecer e conviver com seus jogos de linguagem – ou critérios com os quais constroem e compartilham seus saberes –, a partir da investigação dos usos que este grupo de indivíduos confere aos signos com os quais se comunica – **os discursos** – como ponto de partida da análise. (SALADANHA, 2011, p.63, grifo nosso).

O que vai ao encontro da discussão de Guedes (2010) sobre os desdobramentos do conceito de dialogismo. O autor inicia essa discussão apresentando a visão de Barros (2003), para esta autora:

O dialogismo pode ser desdobrado em dois aspectos: o da interação verbal (entre o locutor e o alocutário do texto) e o da interação entre discursos (ou textos), como prefere Barros (2003). Quanto ao primeiro aspecto, o dialogismo é concebido como o espaço interacional entre o eu e o tu ou entre o eu e o outro, no texto. Em outros termos, o dialogismo se apresenta como produto da relação de alteridade existente entre falantes. O outro aspecto do dialogismo é o diálogo entre os discursos que se instala no interior de cada discurso e o define. Este seria o princípio da interdiscursividade. Entretanto, Fiorin (2006a) defende que não se deve tratar o dialogismo por essa perspectiva, “não há dois tipos de dialogismos, o dialogismo é sempre entre discursos, o interlocutor só existe na qualidade de portador do discurso” (GUEDES, 2010, p. 61 e 62).

O interlocutor pode ser, dentre outras opções, o autor de uma publicação. Sobre o conceito de autor Foucault (1992 apud CAMPOS; VENANCIO, 2006) desenvolve argumentação similar a de Fiorin (2006),

Não se trata de negar o homem ou o autor, mas de substituir um sujeito individual por um sujeito coletivo ou transindividual.[...] o “apagar” do autor era em proveito das formas próprias aos discursos, permitindo descobrir o jogo da função autor, ou seja, ver de que maneira se formou e funcionou este conceito. [...] interessam-lhe as fundações e retomadas da materialidade do conceito de autor (suas funções), independente de qual seja ele. (CAMPOS; VENANCIO, 2006, p. 2).

Os autores também descrevem que Frohmann (2006) considera um conceito restritivo de indivíduo prejudicial aos estudos no campo da ciência da informação argumentando contra um conceito de informação subjetivo e abstrato, pois o foco em um conceito de informação como imaterial o efeito de limitar estudos dos efeitos de regimes de informação às investigações de mudanças na consciência dos indivíduos. O autor conclui que o caráter público e social da informação depende da sua materialidade, visão compartilhada por Campos e Venância (2006, p. 17):

Se, em última instância, a Ciência da Informação remete-se ao indivíduo, não deve fazê-lo de maneira excessivamente direta, delimitando-o e limitando-o a priori. A construção holística e interdisciplinar (eventualmente transdisciplinar) de seu objeto de estudo é capaz de constituir o indivíduo em suas múltiplas dimensões, como um sujeito no mundo e entre documentos, talvez a única maneira viável de atingi-lo.

Percebe-se, assim, uma forte ligação entre e a forma como o sujeito é visto – como um ser social e não individual – e a importância dada, nos estudos da área de informação, aos documentos e à bibliografia.

Frohmann (2009) desenvolve uma abordagem para os conceitos de documento e documentação que tem como objetivo:

Não tanto a precisão e exatidão de uma representação científica do que documentos e documentação podem ser, mas forjando conceitos segundo a visão de Deleuze, buscando aprimorar

o seu poder e força, com foco no que fazem mais do que ao que significam ou representam. (FROHMANN, 2009, p. 301, tradução nossa⁵³).

Nesse sentido, sobre o que se pode fazer com os documentos e o papel que desempenham na nossa sociedade temos o trabalho de Almeida, Cedon e Pinheiro (2012), que com base na teoria dos atos da fala de Austin, apresentam uma proposta para caracterização de documentos que considera a dimensão pragmática, a qual complementar as dimensões sintática e semântica. A conexão entre atos da fala e atos do documento é estabelecida pela noção de intencionalidade⁵⁴.

Consideradas as funções e as práticas, um documento adquire a capacidade de concretizar certos tipos de intencionalidade humana e de ocupar papéis no âmbito das instituições, possibilitando que as ações de incontáveis indivíduos tornem-se coordenadas (SMITH, 2010). De fato, atos do documento, assim como atos da fala, correspondem ao que as pessoas estão aptas a fazer através da intencionalidade (ALMEIDA; CEDON; PINHEIRO, 2012, p. 109).

Os autores apresentam a definição e caracterização do que vem a ser um documento, destacando os seguintes aspectos: “i) documentos implicam em *intencionalidade humana*; ii) um documento pode estar associado a *poderes normativos*; iii) um documento pode instruir, desempenhando papel crucial nas *interações sociais*.” A partir da caracterização do conceito de documento os autores descrevem a

⁵³ Not so much the precision and accuracy of a scientific representation of what documents and documentation might be, but forging concepts in a Deleuzian spirit, seeking to enhance their power and force, with more concern for what they do than for what they mean or represent.

⁵⁴ “O termo *intencionalidade* não diz respeito ao sentido ordinário da palavra “intenção”. O termo em geral é referenciado no âmbito de modelos de recuperação da informação e em Ciência da Informação como um sinônimo para o termo *atinência/aboutness* (FAIRTHORNE, 1969; HJØRLAND, 2001). Entretanto, a busca por uma caracterização complementar às dimensões sintática e semântica demanda definição capaz de integrar a dimensão pragmática. Nesse contexto, *intencionalidade* diz respeito ao poder da mente para se referir a algo, para representar, ou para simbolizar entidades, propriedades e situações; ou para ter a atenção dirigida para certa meta ou coisa (JACOB, s.d).” (ALMEIDA; CEDON; PINHEIRO, 2012, p. 107)

linguagem e o documento como entidades sociais, explicando o mecanismo pelo qual obrigações em sociedade são criadas a partir do uso da linguagem. De modo similar, no entendimento dos autores supracitados, assim como a fala, os documentos produzem efeitos sociais e como documentos são entidades que se mantêm ao longo do tempo, tais efeitos nem sempre são diretamente observáveis, pois podem se manifestar em contextos temporais e/ou geográficos distintos (ALMEIDA; CEDON; PINHEIRO, 2012).

Os documentos, necessariamente, contêm uma determinada linguagem. Hjørland(1998a, p. 21, tradução nossa⁵⁵) ao apresentar as características básicas da teoria dos jogos de linguagem descreve que “palavras em nossa língua só tem sentido na medida em que existem critérios públicos para a sua aplicação. Consequentemente, não pode haver linguagem interna ou privada.”

Randles, Blades e Fadlalla (2012, p. 75, tradução nossa) desenvolvem uma proposta na qual se considera que o conhecimento pode ser dividido por uma espécie de parede,

De um lado desta parede está o conhecimento semântico, que tem uma perspectiva interna e não é bem entendido. No outro lado da parede está o conhecimento pragmático, que tem uma perspectiva externa.

Na abordagem pragmática em CI não se nega a presença de uma estrutura interna, psicológica ou cognitiva, para o desenvolvimento do conhecimento. Só se considera mais útil o estudo da estrutura externa, que está presente na bibliografia e nos diferentes tipos de documentos que registram a comunicação, a troca de informações⁵⁶.

4.1.7 Comunicar é mais importante do que representar

A representação do conhecimento é tradicionalmente vista como o objetivo da OC. Um SOC visa representar o conhecimento de determinado domínio, para organizar a informação e otimizar o processo de RI. O que, na nossa opinião, pode ter relação com a forma como o

⁵⁵ Words in our language have only meaning insofar as public criteria for their application exist. Consequently, there can be no inner or private language.

⁵⁶ Além dos trabalhos citados ao longo dessa seção destacamos Hjørland (2002), Hjørland e Albrechtsen (1995) e Rodrigues e Caricatti (2009) como trabalhos que também contribuíram na formulação do enunciado 6.

conhecimento em si é visto, não como algo construído, mas sim como algo descoberto⁵⁷. A partir da descoberta o homem também representaria o mundo real. Gonzalez de Gomez (1993), no entanto, considera que:

A representação não consiste em uma dimensão necessária da relação gnosiológica do homem com o mundo, mas em um constructo sócio-cultural constituído nas relações de uns homens com outros homens. Na busca de novos pontos de partida, a área dos estudos da organização e da representação do conhecimento – acreditamos – deverá passar pela reconstrução de algumas premissas epistemológicas –subjacentes às mudanças de locus do conhecimento/informação nas modernas formações sociais. (GONZALEZ DE GOMEZ, 1993, p. 2).

As relações entre os homens se dá por meio da comunicação, mesmo que a linguagem não seja a verbal. Rafferty e Hilderley (2007 apud GUEDES, 2010, p. 17), baseados no pensamento de Bakhtin, “sugerem que a indexação deve ser compreendida como uma prática comunicativa”. O que elevaria as ações de organização temática da informação à dimensão reflexiva em que se torna possível estabelecer um diálogo mais consistente com os elementos conceituais presentes nos estudos da linguagem.

Para Novelino (1998) linguagens formalizadas são imprescindíveis para ações de transferência da informação (ou comunicação formal da informação). Essas linguagens refletem uma grande preocupação com aspectos representacionais da linguagem, quando deveriam refletir preocupação maior com o aspecto comunicacional. A partir dessa constatação a autora desenvolve seu trabalho que estabelece fundamentos teóricos, com base no pensamento de Wittgenstein, Austin e Searle, para linguagens formalizadas centradas na comunicação, denominadas linguagens para a transferência de informação. Com relação ao pensamento desses autores a autora considera, resumidamente, que:

⁵⁷ Ver item 4.1.1

O fundamental em Wittgenstein é o princípio do jogo de linguagem. Em Austin, o ato ilocucionário, que é a expressão linguística em seu contexto de produção e de entendimento. Em Searle, a visão de que uma mesma proposição pode ser proferida em diferentes atos ilocucionários e o que se afirma é que usuários diferentes, com intenções de uso diferentes, podem usar expressões linguísticas semelhantes e o que permitirá visualizar essas diferenças serão não as proposições, mas os atos ilocucionários, as conjunturas de realização das proposições. (NOVELINO, 1998, p. 146).

Sobre as diferenças entre as linguagens de representação e as de transferência de informação está o fato de que as de representação consideram exclusivamente a essência de cada informação analisada, enquanto que as linguagens mais centradas na comunicação, as de transferência como a autora propõe, consideram os contextos de produção e de uso da informação, como uma forma de aproximar uma linguagem formalizada da linguagem de uso ordinário. (NOVELINO, 1998).

Saldanha (2008, p.22) considera que:

Ao ir contra o horizonte de uma racionalidade e de um cientificismo estáticos, essencialistas – a representação –, o pragmatismo informacional procura discutir que nenhuma classificação de mundo dura mais que sua linguagem de uso – e que, mesmo entre os becos e travessas que jogam com suas palavras, a dinâmica deste uso transforma permanentemente a estrutura das formas e dos significados. Não há nada definitivo na esfera das relações sociais. Tudo está sob a dinâmica permanente dos contextos com os quais os jogos de linguagem se constituem, se transformam e desaparecem.

Discussão semelhante, sobre o que considerar na OI e OC (a essência ou também aspectos contextuais mais abrangentes), está a diferenciação entre objeto dinâmico e objeto imediato na teoria de Peirce e a relação destes com o potencial de conhecimento dos conceitos, apresentada por Thellefsen e Thellefsen (2004):

A relação entre o conceito e seu potencial de conhecimento é de tal forma que o conceito nunca vai comunicar todo o seu potencial de informações. Da mesma forma, existe uma relação indissolúvel entre a noção de Peirce sobre o objeto imediato e dinâmico. O objeto imediato pode nunca captar plenamente o potencial de conhecimento do objeto dinâmico ou de determinado conceito. O *significance-effect* é uma imagem congelada das informações do conceito exatamente da mesma forma que uma fotografia (por exemplo, um retrato) é o objeto imediato do signo, o objeto que vemos na foto, onde o objeto dinâmico é o indivíduo vivo que em um instante é mantido congelado na imagem. [...] O conhecimento ideal contido nos conceitos é idêntico ao objeto dinâmico. As diferentes representações / interpretações do objeto dinâmico em forma de objetos imediatos correspondem às interpretações dos conceitos. (THELLEFSEN; THELLEFSEN, 2004, p. 183, tradução nossa⁵⁸)

Essas diferentes interpretações são tradicionalmente vistas como um problema. O que é questionado por Novelino (1998, p. 139), pois para a autora “a ambiguidade ou a plurivocidade dos termos, ao invés de dificultar a comunicação ou, no caso, a transferência da informação, tornará possível agregar ao termo o sentido que é dado pela conjuntura na qual ele se insere: o texto e o contexto de produção do texto”.

Complementando essa afirmação apresentamos a visão de Frohmann (2009) que, apesar de não se referir aos SOC especificamente, é pertinente:

⁵⁸ The relation between the concept and its potential knowledge is in such a way that the concept will never communicate all of its potential information. In the same way, there exists an insoluble relation between Peirce's notion on the immediate and dynamical object. The immediate object can never fully capture the knowledge potential of the dynamical object or a certain concept. The significance-effect is a frozen picture of the information value of the concept exactly in the same way a photograph (i.e. a portrait) is the immediate object of the sign, namely the object we see in the picture, where the dynamical object is the living individual who in an instant is maintained frozen in the picture. [...] The ideal knowledge contained by the concepts is identical with the dynamical object. The different representations/interpretations of the dynamical object in the shape of immediate objects correspond to the member's interpretations of the concepts.

Pode faltar invariabilidade aos nossos conceitos [...], as considerações sobre como seguir uma regra podem ser múltiplas, mas ainda assim podemos nos comunicar e falar de forma significativa sobre documentos ou qualquer outra coisa, porque não há nenhuma ambiguidade geral, inerente à linguagem, só específicas, ambiguidades subordinadas a falantes específicos em condições específicas. (FROHMANN, 2009, p. 294, tradução nossa⁵⁹).

Para Frohmann (2009):

Os critérios, regras e definições podem ser especificados (e executados) para fins particulares, mas nem sempre já estão lá, guiando-nos sobre como falamos, e nós geralmente não somos obrigados a fornecê-los, nem há qualquer ambiguidade generalizada da linguagem que representa uma ameaça ao pensamento e à comunicação como tal – mesmo quando imprecisão e ambiguidade pode representar uma ameaça para tipos específicos de conversa que dependem de definições bem formadas, significados fixos e regras claras. (FROHMANN, 2009, p. 295, tradução nossa⁶⁰)

O autor faz essa afirmação no contexto de um artigo no qual discute os conceitos de “documento” e “documentação” sem apresentar novas definições, mas argumentando ou descrevendo o que se faz com documentos. Justifica essa opção com base nas observações de Putman (2001) sobre como revoluções científicas, como a teoria geral da relatividade de Einstein e a geometria não-euclidiana, ocorreram e novas

⁵⁹ Our concepts can lack fixity (and to this degree Wittgenstein is Derridean), what counts as following a rule can be multiple, yet we still can communicate and speak meaningfully, about documents and anything else, because there is no general, inherent ambiguity of language, only specific, contingent ambiguities for particular speakers under specific conditions.

⁶⁰ Criteria, rules and definitions can be specified (and enforced) for particular purposes, but are not always already there, guiding us as we speak, and we are not generally obliged to provide them, nor is there any generalized ambiguity of language posing a threat to thought and communication as such – even when imprecision and ambiguity might pose a threat to specific kinds of talk that rely upon well-formed definitions, fixed meanings, and clear rules.

teorias foram aceitas ao em vez de serem consideradas irracionais. Sobre esse aspecto, Frohman (2009, p. 296, tradução nossa) exemplifica que: “certamente, o primeiro físico que aceitou a Teoria Geral da relatividade foi considerado como racional, na verdade como sendo dotado de um alto grau de percepção científica”, mas tal aceitação não poderia ser vista como um tipo de regularidade, muito menos como uma regra”. Para exemplificar que, mesmo quando tratamos de conceitos científicos, mudanças conceituais não se constituem em arbitrariedade, mas, ao contrário, manifestam a mais alta ordem da racionalidade científica o autor cita conceitos da Física, como “linha reta” e, “simultaneidade” que sofreram mudanças com Einstein. Para concluir essa perspectiva, lembra que Putnam destaca que “Einstein contou a sua história e mostrou como aplicá-la”.

Assim, o autor discorre sobre o conceito de documento, argumentação que pode ser considerada para conceitos de maneira geral:

Ao se projetar a palavra “documento” a novas situações não precisamos fazê-lo com base em novos critérios, regras, significados ou definições. Às vezes, outros irão dialogar, ou “continuar” a partir de nossas extensões; em outros momentos, continuarão da forma como eles concebem, mas que nós não fomos capazes de prever. Alguns podem recusar-se a ir em nossa direção, podem não achar relevante para as suas finalidades fazê-lo. Mas se queremos estender o conceito de documentação a novas situações – se queremos multiplicar o conceito de documento – podemos (como Einstein) contar uma história e mostrar como aplicá-lo e, assim, conectar nossas extensões de teorias, argumentos e recursos intelectuais relevantes para tentar trazer uma compreensão ampliada do que documentos e documentação podem ser. (FROHMANN, 2009, p. 296, tradução nossa⁶¹).

⁶¹ If we project the word “document” onto new situations we need not do so on the basis of new criteria, rules, meanings, or definitions. Sometimes others will chime in, or will “go on” from our extensions as we do; at other times, in ways they see as flowing from them but that we did not anticipate. Some may refuse to go in our direction, or not find it relevant to their purposes to do so. But if we want to extend documentation to new situations – if we want to multiply the concept of the document – we can (like Einstein) tell a story and how to apply it, and connect our extensions to theories, arguments and

Ainda que o autor não trate especificamente dos conceitos em SOC, mas sim da construção de teorias, suas observações contribuem para nossa argumentação de que a comunicação é uma finalidade mais adequada do que a representação nesses sistemas.

A comunicação exige uma linguagem, que nomeia as coisas, para facilitar o uso dos objetos, como afirma Hjørland (2003). A indexação é a nomeação dos assuntos de um documento, daí a responsabilidade desse importante processo na OI que deve facilitar a utilização de assuntos. Mas, a seleção dos assuntos, depende da análise de relevância e qualidade da informação e ao nomear um assunto na indexação e deixar de nomear outro, aquele que foi excluído tem seu uso dificultado. Hjørland (1992 apud HJØRLAND, 2003, p.98, tradução nossa⁶²) define o assunto de um documento como a “potencialidade epistemológica (ou informativa) do documento. Um documento tem determinadas potencialidades, e o trabalho dos índices (e outros SAP⁶³) é identificar essas potencialidades.” No item 4.1.8 a seguir continuamos a reflexão sobre esse aspecto⁶⁴.

4.1.8 Guiar para uma escolha informada

Uma questão que nos chama a atenção ao analisar o pensamento dos autores que tratam sobre a pragmática na CI é a crítica aos atuais modelos de recuperação da informação. Blair (2003) considera que:

[...] pode ser extremamente difícil conceber sistemas de informação radicalmente diferentes ou melhorados, porque estamos praticamente bloqueados na maneira de pensar sobre a recuperação da informação, que se materializa

relevant intellectual resources to try to bring about an enlarged understanding of what documents and documentation can be.

⁶² As the document's epistemological (or informative) *potentialities*. A document has given potentialities, and the job for indexers (and other SAPs) is to identify those potentialities.

⁶³ Pontos de acesso de assuntos.

⁶⁴ Além dos trabalhos citados ao longo dessa seção destacamos Almeida (2012b), Gonzalez de Gomez (1995), Gracioso (2010 a; 2012), Hjørland (2011d; 2008b) e Saldanha (2010) como trabalhos que também contribuíram na formulação do enunciado 7.

pelos sistemas existentes. (BLAIR, 2003, p.14, tradução nossa⁶⁵).

O processo de RI, muitas vezes, é visto como aquele no qual o investigador tem algo em mente, a necessidade de informação, que é traduzida em uma consulta de pesquisa. No entanto, com base nas afirmações de Wittgenstein, o autor explica que a forma como se pensa a necessidade de informação é condicionada pela linguagem de recuperação disponível. Na medida em que essa linguagem é limitada, assim será o pensamento sobre o que se quer, sobre a necessidade de informação (BLAIR, 2003).

A avaliação dos sistemas de recuperação de informação (SRI) leva em conta, principalmente e talvez unicamente, os índices de precisão e revocação. Esses índices são calculados a partir do julgamento de relevância, entendida tradicionalmente como o que o usuário já teria em mente, aquilo que o usuário quer encontrar com a busca.

Pensar na diferença entre querer e precisar pode ajudar a entender essa questão. Aquilo que o usuário quer, que tem relação com o conceito tradicional de relevância, pode não ser o que ele precisa. O precisar tem uma ligação direta a aspectos funcionais e práticos enquanto querer é algo mais subjetivo. Assim, o precisar nos parece ser compatível com a definição de relevância, de Hjørland e Christensen (2002, p. 964, tradução nossa⁶⁶), “Algo (A) é relevante para uma tarefa (T) se aumentar a probabilidade de realizar o objetivo (G), que está implícito em T”.

Frohmann (1990, p. 98, tradução nossa⁶⁷) também aborda esse tema e apresenta a seguinte questão: “A recuperação de textos deve satisfazer uma necessidade ou satisfazer um desejo?”.

Desejos são explicitamente reconhecidos e admitidos; eles refletem os objetivos dos agentes, os propósitos e intenções. Nem todas as necessidades são conhecidas, [...] Por exemplo, nem todos sabem o que precisam para evitar AIDS, e nem todo mundo deseja o que precisa. A

⁶⁵ May be extremely difficult to design radically different or improved retrieval systems, because we are virtually locked into the way of thinking about retrieval that is embodied by existing systems.

⁶⁶ *Something (A) is relevant to a task (T) if it increases the likelihood of accomplishing the goal (G), which is implied by T.*

⁶⁷ Does text retrieval fulfil a need, or does it satisfy a want?

identificação das necessidades depende de uma concepção da natureza humana e do mundo social; desejos podem ser identificado por meio de um questionário. Se apenas a satisfação do desejo for considerada a finalidade da recuperação de textos, então a maioria das regras de indexação para as práticas de recuperação servirão à forma de organização social predominante. Entre as regras de indexação importantes para a satisfação do desejo no capitalismo de consumo, por exemplo, são aquelas que representam eficientemente bens para o consumo. Por outro lado, se a recuperação de textos deve satisfazer as necessidades, então as regras para a sua prática não só podem ser inconsistentes com os objetivos da ordem social dominante, mas também ser antagonistas a eles. (FROHMANN, 1990, p. 98, tradução nossa⁶⁸).

Essas e outras questões exigiriam uma análise bem mais ampla, ou uma análise política como o autor sugere. Em todo caso, o que essa questão deixa claro é que as “diferentes concepções sobre o papel social da recuperação de textos determinarão os tipos de regras de indexação que construímos”. (FROHMANN, 1990, p. 98, tradução nossa⁶⁹).

Basear-se prioritariamente nos “desejos” e naquilo que o usuário já tem em mente, o que ele quer encontrar, pode ser limitativo e acabar ocultando pontos de vistas importantes presentes na literatura. Esse entendimento é subsidiado por Hjørland (2007b, p. 392, tradução nossa⁷⁰),

⁶⁸ Wants are explicitly recognised and admitted; they reflect the agents' goals, purposes, and intentions. Not all needs are known, and some fly in the face of wants. For example, not everyone knows what they need to prevent AIDS, and not everyone wants what they need. Identification of needs depends upon a conception of human nature and the social world; wants can be identified by questionnaire. If want satisfaction alone is considered the end of text retrieval, then most indexing rules will serve the retrieval practices of the prevailing form of social organisation. Among the important indexing rules for want satisfaction in consumer capitalism, for example, are those which efficiently represent goods for consumption. On the other hand, if text retrieval is taken to fulfil needs, then rules for its practice may not only be inconsistent with the aims of the dominant social order, but may also be antagonistic to them.

⁶⁹ Differing conceptions of the social role of text retrieval will determine the kinds of indexing rules we construct.

⁷⁰ In IR, the basic function of semantic relations is to contribute to the increase of recall and precision. For example, the inclusion of synonyms and broader terms in a query may

Na RI a função básica das relações semânticas é contribuir para a otimização da precisão e da revocação. Por exemplo, a inclusão de sinônimos e termos mais abrangentes em uma consulta pode contribuir para o aumento da revocação, enquanto que a diferenciação de homônimos e a especificação dos termos pode aumentar a precisão. Desse modo, a ampla utilização das relações semânticas padrão empregadas em tesauros pode ser explicada funcionalmente.

Ao avaliar a RI apenas pelos índices de precisão e revocação, com base na ideia tradicional de relevância, pode-se considerar que as relações semânticas utilizadas nos tesauros seriam suficientes. No entanto, verifica-se na literatura da área o incentivo ao uso de outros tipos de relações semânticas, o que leva Hjørland (2007b, p. 393, grifo nosso, tradução nossa⁷¹) ao seguinte questionamento: “Como devemos explicar essa demanda por um conjunto muito mais rico de relações do que as normalmente usadas em tesauros?”, e esclarece que “a resposta pode implicar em uma crítica aos **tradicionais** índices de precisão e revocação como forma de compreender a RI”. Uma função adicional que as relações semânticas em um SOC poderiam desempenhar seria explicitar os pontos de vista distintos presentes nos domínios do conhecimento (HJØRLAND, 2007b).

De acordo com Gracioso (2010b, p. 141) “é importante que pesquisadores (buscadores de informações) tenham a dimensão da indeterminação da relação entre o que eles descreverão e o que, diante do que foi recuperado, terão de discriminar”. Para isso, conforme afirma Hjørland (2007b, p. 390, tradução nossa⁷²) “talvez a tarefa mais importante do profissional da informação é fazer os diferentes interesses e paradigmas visíveis para que o usuário possa fazer uma escolha informado.” Assim, talvez perceber que aquilo que ele esperava encontrar a princípio, aquilo que ele queria, não era o que ele efetivamente precisava.

contribute to increased recall, whereas the differentiation of homonyms and the specification of terms may increase precision. In this way, the wide use of the standard semantic relations employed in thesauri may be explained functionally.

⁷¹ How should we explain this demand for a much richer set of relations than that ordinarily used in, for example, thesauri? The answer may imply a criticism of the traditional recall/precision way of understanding IR.

⁷² Perhaps the most important task of the information professional is to make the different interests and paradigms visible so that the user can make an informed choice.

Para Hjørland(1998a, p. 29):

O que os usuários precisam não são seleções "neutras" dos documentos. Relevância não é uma escala unidimensional que se baseia em propriedades quantitativas. Os usuários precisam de "mapas" de estruturas de informação, o que pode ajudá-los a se orientar e aperfeiçoar os seus argumentos de busca. Esses mapas devem refletir as abordagens básicas e devem mostrar os significados mais ou menos ocultos, interesses e objetivos dos documentos. (HJØRLAND, 1998a, p.28, tradução nossa⁷³).

Fainburg (2009) tem como objetivo explicar como a necessidade de informação do usuário poderia ser comparada com a aprendizagem, na qual o objetivo de buscar informações é encontrar informações relevantes para o cumprimento de um objetivo específico. Para tanto apresenta e compara o modelo *information seeking process* (ISP) de Carol Kuhlthau e o modelo de resolução de problemas de John Dewey (1859-1952), concluindo que:

Enquanto Kuhlthau investiga o usuário individual, a filosofia educacional de Dewey e o pragmatismo clássico descobrem que a cooperação coletiva é importante. Também parece importante integrar o processo de busca de informações e resolução de problemas com as teorias de como escrever artigos acadêmicos e dissertações, como proposto por Andersen (2006) e Rienecker e Stray Jørgensen (2005). Isso sugere que, se os usuários de sistemas de informação passarem com sucesso pelas fases do modelo de Kuhlthau e do modelo de resolução de problemas de Dewey, eles estarão se tornando usuários com competência informacional ou profissionais reflexivos, de

⁷³ What the users need are not "neutral" selections of the documents. Relevance is not a one-dimensional scale based on quantitative properties. Users need "maps" of information structures, which can help them to be oriented and to refine their search arguments. Such maps should reflect the basic approaches and should uncover the more or less hidden meanings, interests and goals in documents.

acordo com Schön (1987). (FAINBURG, 2009, p.464, tradução nossa⁷⁴).

O profissional reflexivo ao qual o autor se refere tem relação com o pensamento reflexivo também desenvolvido por John Dewey. Dewey (1979 a, apud GASQUE, 2008, p. 84):

Faz distinção entre o pensamento reflexivo, aquele com intenção de descobrir a verdade, mesmo que provisória, - dos outros processos mentais, algumas vezes denominados pensamentos, mas que não pretendem a verdade.

Guasque (2008, p. 77) considera que:

O pensamento reflexivo pode ser utilizado como estratégia cognitiva na construção das competências necessárias à busca e ao uso da informação, possibilitando a compreensão mais profunda das questões, fenômenos e processos envolvidos por meio da percepção das relações de identificação dos elementos, da análise e interpretação dos sentidos.

Para a autora o conceito de pensamento reflexivo de Dewey seria melhor entendido a partir da análise de sua vida e obra, nesta análise afirma-se que “a filosofia Deweyana compreende a pessoa livre para escolher as ações e tomar decisões em um processo social, democrático, de supostas oportunidades iguais para todos.” (GASQUE, 2008, p. 82).

Fainburg (2009) descreve que Dewey é considerado o fundador de conceitos como “aprender fazendo” e “aprendizagem ao longo da vida”, os quais considera essenciais na sociedade atual em que o conhecimento tem se tornado cada vez mais efêmero. De acordo com o autor, Dewey era contra qualquer tipo de dualismo, por exemplo, entre o pensamento e a ação ou a teoria e a prática. “Pensamento sem ação é

⁷⁴ While Kuhlthau investigates the individual user, Dewey's educational philosophy and the classical pragmatism finds that collective cooperation is important. It also seems important to integrate the process of information seeking and problem solving with theories of how to write academic papers and dissertations as proposed by Andersen (2006) and Rienecker and Stray Jørgensen (2005). This suggests that if users of information systems have passed successfully through Kuhlthau's ISP phases and Dewey's problem solving model, they are becoming information literate users or reflective practitioners according to Schön (1987).

percebido como estéril e ação sem pensamento como rotina, enquanto que o ideal é a *'thoughtful action'*.” (FAINBURG, 2009, p. 462, tradução nossa⁷⁵).

Ao se compreender que uma necessidade de informação tem relação com a realização de uma tarefa, o cumprimento de um objetivo, se percebe que essa necessidade pode ultrapassar aquilo que se desejava ou se esperava encontrar em uma busca por informação. Nesse sentido o processo de RI pode ser comparado com a aprendizagem, como sugere Fainburg (2009), e o conceito de relevância é alterado⁷⁶.

4.1.9 Relações semânticas para contextualizar os pontos de vista distintos

Com base nas concepções de Wittgenstein, sobre o papel da prática com a linguagem para a formação do significado, Frohmann (2004, p. 396, tradução nossa⁷⁷) considera que:

Também devemos mudar a nossa atenção de imagens mentalistas de informação para práticas com documentos. A perspectiva de Wittgenstein anuncia uma mudança de teorias da informação para descrições de práticas documentais.

O autor afirma também que “a informatividade de um documento depende de certos tipos de práticas, e porque a informação emerge como um efeito de tais práticas, as práticas documentais são ontologicamente primárias à informação”. (FROHMANN, 2004, p. 387, tradução nossa⁷⁸).

Nesse sentido, é interessante notar também que o estabelecimento de relações semânticas pode ser observado nas práticas de RI. Conforme Hjørland (2007b) descreve é importante ter em mente que conceitos e relações semânticas são usados, por exemplo, para a expansão de

⁷⁵ Thoughts without action are perceived as barren and action without thinking as routine, while the ideal is “thoughtful action” (Dewey, 1916).

⁷⁶ Além dos trabalhos citados ao longo dessa seção destacamos Cavaleri (2008), Diefenbach e Hassenzahl (2011) e Hjørland (2002; 2011c; 1998b) como trabalhos que também contribuíram na formulação do enunciado 8.

⁷⁷ Too does it shift our attention away from mentalistic pictures of information and toward practices with documents.

⁷⁸ The informativeness of a document depends on certain kinds of practices with it, and because information emerges as an effect of such practices, documentary practices are ontologically primary to information.

consulta, bem como para a precisão. Desta forma, uma “necessidade de informação” específica determina quais relações são fecundas e quais não são. Uma relação semântica que aumenta a precisão e a revocação em uma determinada pesquisa é relevante para aquela situação específica. Pesquisadores de informações criativos fornecem estratégias de busca que recuperam um conjunto fecundo de documentos, combinando termos de formas incomuns. Termos diferentes podem ser combinados usando o operador booleano OR em determinada pesquisa, por implicação, em algumas situações eles são considerados como termos equivalentes (ou sinônimos), mesmo que normalmente não sejam considerados sinônimos. Antônimos, por exemplo, em RI, podem ser úteis vezes porque certos fenômenos podem ser discutidos em relação aos seus opostos. Pode-se afirmar, portanto, que em certos casos antônimos e sinônimos desempenham a mesma função na RI e que a relação semântica entre termos sinônimos, a de equivalência, poderia ocorrer entre termos antônimos.

Relações semânticas dizem respeito a determinada tarefa ou situação e nem todos os usuários compartilham sempre a mesma visão de que determinados termos são equivalentes. Por outro lado, não é possível se apoiar numa teoria semântica que tem como base uma visão individualista/idiossincrática de conceitos e semântica para mais de um usuário ou situação. Precisamos de princípios mais estáveis para determinar as relações semânticas. Precisamos de uma teoria semântica sobre o significado de palavras como formas de práticas tipificadas. O conhecimento sobre a semântica das práticas tipificadas pode então ser utilizado por pesquisadores de informação, a fim de incluir ou excluir determinados documentos. (HJØRLAND, 2007b, p. 381, tradução nossa⁷⁹).

⁷⁹ Semantic relations relate to a given task or situation and not all users of a given set of semantic relations will share the same view of which terms are equivalent. On the other hand, it is clear that if we base a semantic theory on an individualistidiosyncratic view of concepts and semantics, it is not possible to design systems for more than one user or situation-an absurd conclusion. We need more stable principles on which to determine semantic relations. We need a semantic theory about the meaning of words as forms of *typified practices*. Knowledge about semantics in typified practices may then be used by information searchers in order to include or exclude certain documents.

Nesse sentido, Hjørland(2003, p. 97, tradução nossa⁸⁰), aponta que:

As pessoas que conhecem muito bem determinadas bases de dados, inclusive a literatura sobre o assunto coberto e sua linguagem especializada, provavelmente empregam este tipo de conhecimento de forma tácita. O trabalho para estudos da informação é ajudar a explicar os princípios subjacentes. Assim como pessoas comuns podem falar uma língua e usar a gramática, *experts* em linguística têm de explicar o que os usuários competentes fazem.

A explicação desses princípios subjacentes, por meio de relações semânticas por exemplo, exigiria então o conhecimento não só das ferramentas da base de dados, mas também sobre o assunto.

Você não é capaz de determinar as relações semânticas entre as palavras “Copenhague” e “Dinamarca”, a menos que você saiba que Copenhague é uma parte da Dinamarca. (HJØRLAND, 2007b, p.375 e 376, tradução nossa⁸¹).

Nesse sentido, muitas vezes se recomenda que um SOC seja desenvolvido em equipe e que conte com especialistas na área que o sistema cobre. No entanto, Hjørland (2007 b, p. 390, tradução nossa⁸²) afirma que, “os cientistas da informação podem sentir-se mais seguro se contarem com “especialistas” em vez de documentos, mas documentos relevantes são escritos por especialistas e são fontes igualmente válidas”.

⁸⁰ People, who know certain databases very well, including the subject literature and subject language they cover, probably employ this kind of knowledge in tacit ways. The job for information studies is to help to explicate the underlying principles. Just as ordinary people can speak a language and use grammar, linguistic experts have to explicate what competent users do.

⁸¹ You cannot determine the semantic relations between the words “Copenhagen” and “Denmark” unless you know that Copenhagen is a part of Denmark.

⁸² Information scientists may feel safer if they rely on “experts” rather than documents, but relevant documents are written by experts and are equally valid sources, if not more so.

Conhecimento do domínio é um problema não só para a CI, mas também para a linguística e muitas meta-ciências (como a ciência cognitiva e a sociologia da ciência). Teorias cognitivas e linguísticas sobre conceitos, significado e semântica é fortemente restringida por tentativas de evitar o “mundo do conhecimento”. A importância do conhecimento do assunto tem implicações teóricas sobre como os conceitos devem ser definidos e as relações semânticas determinadas (seja por humanos ou por máquina). Tem implicações para responder a pergunta: Que tipo de informação é necessária, a fim de determinar as relações semânticas entre dois termos A e B? (HJØRLAND, 2007b, p. 377, tradução nossa⁸³).

De acordo com o autor, teorias semânticas, usadas na OC, devem fornecer implicações metodológicas para determinar significados e relações semânticas. No entanto, muitas vezes tais implicações não são claras, o que torna as teorias vagas e inúteis. Frohmann (1983 apud HJØRLAND, 2007b) reconhece que os problemas na classificação devem ser vistos como problemas relacionados às teorias semânticas. Ele observa que conceitos podem ser agrupados ou classificados de diferentes maneiras, de acordo com o princípio de divisão adotado. Apresenta duas teorias semânticas: uma sustenta que as categorias às quais um conceito pertence são dadas *a priori* como parte do significado do termo para esse conceito; a outra defende que as categorias às quais pertencem um conceito devem ser encontradas na literatura específica ou no discurso e, conseqüentemente, as relações semânticas não são dadas *a priori*, mas são formulados *a posteriori*.

Conforme já tratamos na fundamentação teórica desse trabalho, quando nos referimos às relações paradigmáticas (*a priori*) e às sintagmáticas (*a posteriori*). as relações de equivalência, sinonímia, e as hierárquicas são tradicionalmente consideradas relações *a priori*. No

⁸³ Domain knowledge is a problem not only for IS but also for linguistics and many metasciences (such as cognitive science and the sociology of science). Much cognitive and linguistic theory regarding concepts, meaning, and semantics is strongly constrained by attempts to avoid “world knowledge.” The importance of subject knowledge has theoretical implications for how concepts should be defined and semantic relations determined (whether by human or by machine). It has implications for answering the question: What kind of information is needed in order to determine the semantic relations between two terms A and B?

entanto, sinônimos puros ou perfeitos são raros, Hjørland e Pedersen (2005, p. 586, tradução nossa⁸⁴) esclarecem que:

Termos podem ser considerados como mais ou menos sinônimos ou homônimos (ou não relacionados) dependendo das perspectivas (Frohmann, 1983). Portanto, essa propriedade das descrições também pode ser carregada de teoria.

Assim, os autores explicam a abordagem que desenvolveram, que é também um bom exemplo, na nossa opinião, da proposta de Hjørland (2003, 2007 b). Por isso a descrevemos nessa longa citação a seguir:

Os problemas de sinônimos e homônimos estão relacionados com o problema de classificação porque a classificação é supostamente para unir termos diferentes com o mesmo significado e para distinguir diferentes significados de um mesmo termo. Spärck Jones tem investigado algoritmos que podem ser utilizados para classificar automaticamente documentos com base na identificação de sinônimos e homônimos. Nós sugerimos o que acreditamos ser uma nova abordagem para este problema. Uma abordagem que compartilha algumas características com o uso de coleções de teste (Spärck Jones e Van Rijsbergen, 1976), mas que é mais interpretativa e de certa forma coloca as coisas de cabeça para baixo. A abordagem que sugerimos baseia-se no estudo das literaturas específicas, evitando qualquer generalização (pelo menos como ponto de partida). [...] A idéia básica é que por esse meio somos capazes de estimar o valor de diferentes discriminadores sem depender de coisas como o julgamento de relevância associados aos experimentos tradicionais em RI. Isso faz com que seja possível construir um algoritmo para a solução dos homógrafos, que, no entanto, não é apenas mais um algoritmo entre outros, mas que é o algoritmo teoricamente ideal, de acordo com a

⁸⁴ Terms may be regarded as more or less synonymous or homonymous (or otherwise related) depending on perspectives (Frohmann, 1983). Hence, this property of the descriptions may also be theory-laden.

literatura específica e a identificação manual de homógrafos. Mas por que preferimos essa abordagem quando não acreditamos que é generalizável a outros homógrafos ou outras literaturas? A resposta a essa pergunta é que acreditamos que nós podemos gradualmente aprender mais e mais sobre as literaturas específicas e sobre as práticas relativamente padronizados que os autores realizam. Em outras palavras, podemos aprender sobre diferentes maneiras de expressar-se em documentos e sobre padrões relativamente estáveis que podem ser explorados em RI. Isto é completamente diferente de um método positivista de realizar um experimento após o outro, na tentativa de identificar um algoritmo ótimo para determinado fim ou contexto não específicos. (HJØRLAND; PEDERSEN, 2005, p. 588 e 589, tradução nossa⁸⁵).

⁸⁵ The problems of synonyms and homonyms are related to the problem of classification because a classification is supposed to unite different terms with the same meaning and to differentiate between different meanings of the same term. Spärck Jones has investigated algorithms that may be used to automatically classify document representations based on the identification of synonyms and homonyms. We here suggest what we believe to be a new approach to this problem. An approach that shares some characteristics with the use of test collections (Spärck Jones and van Rijsbergen, 1976) but which is more interpretative and in a way puts things upside down[8]. The approach we suggest is based on the study of specific literatures, avoiding any generalization (at least as a starting point). [...]The basic idea is that we hereby are able to estimate the value of different discriminators without depending on such things as relevance judgment associated with traditional IR experiments. This makes it possible to construe an algorithm for the solution of the homograph, which, however, is not just one more algorithm among others, but which is the theoretically optimal algorithm given the specific literature and the manually identified homograph. But why prefer such an approach when we do not believe that it is generalizable to other homographs or other literatures? The answer to this question is that we believe that we may gradually learn more and more about specific literatures and about the relatively standardized practices that authors perform (i.e. about genres). In other words, we learn about different ways of expressing oneself in documents and relatively stable patterns[9] which may be exploited in IR. This is quite different from a positivist method of performing one experiment after another, attempting to identify an optimal algorithm for a non-specific purpose or context.

Hjørland (2003) argumenta que:

Relações semânticas não podem ser estabelecidas essencialmente por premissas universalistas, mas devem ser entendidas principalmente como específicas de um domínio, reveladas (e construídas) pelas disciplinas científicas. O campo de OC na CI não pode ignorar conceitos, teorias e descobertas em disciplinas específicas, e os métodos de OC em CI estão em um nível mais profundo baseados nos mesmos pressupostos filosóficos que os métodos acadêmicos e científicos. Isso implica que a discussão fundamental das bases da OC na CI está fortemente ligada à discussão de diferentes teorias da epistemologia. Estudos epistemológicos têm sido raros na CI, e parece urgente para o nosso campo para atualizar nossas qualificações nessa área. (HJØRLAND, 2003, p. 107, tradução nossa⁸⁶).

Hjørland (2003) explica a influência dos diferentes pontos de vista no estabelecimento das relações semânticas por meio da descrição da evolução do conhecimento científico. Apresenta como exemplo a classificação dos animais. As baleias vivem na água e podem ser classificadas como animais aquáticos, são mamíferos e não peixes. A classificação exige que se encontrem as propriedades semelhantes entre os itens a serem classificados para então agrupá-los. Essas propriedades semelhantes também podem estabelecer outras relações entre os itens. Hjørland (2003, p. 102, tradução nossa⁸⁷) aponta que “a história de todas as ciências naturais documenta a descoberta de que certas entidades que partilham propriedades imediatas, no entanto, pertencem a diferentes espécies.”

⁸⁶ semantic relations cannot primarily be established by universalistic assumptions, but much primarily be understood as domain specific, as uncovered by (and constructed by) scientific disciplines. KO in LIS cannot ignore concepts, theories and findings in specific disciplines, and the methods of KO in LIS are at the deepest level based on the same philosophical assumptions as the methods of science and scholarship. This implies that the fundamental discussion of the basis of KO in LIS is strongly connected to the discussion of different theories of epistemology. Epistemological studies have been rare in LIS, and it seems urgent for our field to upgrade our qualifications in this area.

⁸⁷ The history of all natural sciences documents the discovery that certain entities that share immediate properties nonetheless belong to different kinds.

Assim, o autor afirma que;

Esta realidade mais profunda de espécies e conceitos que a ciência descobre tem implicações importantes para a metodologia de OC. Em primeiro lugar ela desafia muitas abordagens orientadas ao usuário e empiristas. Na medida em que esta visão é correta, as relações entre dois conceitos são, portanto, em relação aos sistemas teóricos (ou paradigmas) em que estão incorporadas. [...] Podemos concluir que as unidades básicas na OC, as relações semânticas entre dois conceitos, devem ser relativas à perspectiva e à teoria da qual elas são considerados. Devido a este fato, OC não pode ser feita apenas a partir de combinações sucessivas de elementos mas deve refletir perspectivas e teorias mais amplas. (HJØRLAND, 2003, p. 102 e 103, tradução nossa⁸⁸).

Inclusive a forma como o autor entende os conceitos, que junto com as relações semânticas formam os SOC, vai ao encontro dessa afirmação. Conceitos, pelo menos para fins de OC, são entendidos como significados socialmente negociados e uma justificativa para esse entendimento, que nos parece oposto ao entendimento do conceito como unidade do pensamento, seria o fato de que:

Se o conhecimento é definido como crença verdadeira justificada (como na tradição platônica), então o conhecimento real é difícil ou impossível de identificar e classificar. É mais proveitoso falar de reivindicações de conhecimento, no lugar do conhecimento em si. Falar em reivindicações de conhecimento sobre as coisas representadas na literatura e as coisas a

⁸⁸ It demonstrates the deeper reality of kinds and concepts that science discovers. It has important implications for the methodology of KO. First and foremost it challenges many user-oriented and empiristic approaches. To the degree that this view is correct, the relations between two concepts are thus relative to the theoretical systems (or paradigms) in which they are embedded. [...] We may conclude that the basic units in KO, the semantic relations between two concepts, may be relative to the perspective and the theory from which they are considered. Because of this fact, KO cannot be done just from successive combinations of elements, but must reflect broader perspectives and theories.

serem classificadas é uma maneira mais cuidadosa de falar, e não há perda real por este modo de dizer. (HJØRLAND, 2003, p. 100, tradução nossa⁸⁹).

Ao nosso ver, isso também justifica o fato de as relações semânticas serem consideradas pelo autor como a unidade básica na OC, pois, na negociação dos significados, que podemos identificar nas reivindicações de conhecimento, formam-se relações semânticas e, ainda que a negociação não termine em um acordo (sem chegar a um conceito unívoco), as relações semânticas estarão lá, e se forem devidamente identificadas ajudarão a comunicar, ou mostrar, para o usuário o contexto, ou os contextos, nos quais o conceito está inserido.

Quando os profissionais de informação classificam documentos, os significados relevantes e propriedades só estão disponíveis com base em algumas descrições. Essa consideração opõe-se à suposição implícita predominante de que todas as propriedades relevantes dos objetos são óbvias para os especialistas em informação, os quais acompanhariam certos princípios estabelecidos e forneceriam uma classificação melhor: objetiva, neutra e universal, daí, tecnicamente eficiente (HJØRLAND, 2003).

Desse modo, assim como as distintas áreas do conhecimento não são neutras e apresentam apenas uma parte da totalidade de descrições possíveis sobre determinado tema, os SOC também não serão.

Não é possível ser neutro, mas é absolutamente inaceitável ocultar visões diferentes e suprimir a capacidade dos usuários para desenvolver seus próprios pontos de vista. (HJØRLAND; PEDERSEN, 2005, p. 593, tradução nossa⁹⁰).

4.1.10 Síntese conclusiva

Nesta seção apresentamos as principais características da abordagem pragmática na OC, campo no qual o estudo do significado,

⁸⁹ If knowledge is defined as justified true belief (as in the Platonic tradition) then real knowledge is hard or impossible to identify and to classify. It is more fruitful to speak of knowledge claims, rather than knowledge itself. To speak of *knowledge claims* as the things represented in the literature and the thing to be classified is a more careful way of speaking, and there is no real loss by this way of speaking.

⁹⁰ It is not possible to be neutral, but is absolutely unacceptable to hide different views and to suppress the users' ability to develop their own points of view.

ou semântica, é de suma importância, o que, de certa forma, nos faz entender a pragmática como um paradigma em semântica, conforme sugere Peregrin (2004; 1999), visão que também é aceita por Birger Hjørland.

A partir dos resultados obtidos podemos dizer que na abordagem pragmática na OC a informação é vista como um fenômeno que é construído. Essa construção se dá em circunstâncias diversas e envolve diferentes atores que podem ter pontos de vista e necessidades distintas, devido aos diversos aspectos contextuais envolvidos na produção e uso da informação e do conhecimento. Assim, a avaliação da qualidade da informação deve levar em conta esses pontos de vista distintos. Tanto a informação quanto a sua qualidade não são vistos como fenômenos constantes e avaliar uma fonte de informação a partir de premissas de verdade ou falsidade é visto como perigoso, pois se deve avaliá-la em relação à forma como aborda as controvérsias e diferentes pontos de vista. Assim como as enciclopédias, e outras fontes de informação do gênero os SOC também devem apresentar os diferentes pontos de vista encontrados na literatura de uma área do conhecimento.

Considera-se mais importante possibilitar uma comunicação efetiva pela compreensão dos diferentes pontos de vista do que a escolha de um como sendo o melhor para representar uma realidade que é, por natureza, heterogênea. As relações semânticas poderiam desempenhar papel crucial nessa comunicação, ao guiar o usuário a partir do seu ponto de vista inicial até aqueles defendidos pela frente de pesquisa sobre o tema.

Na seção a seguir apresentamos a comparação, pontos compatíveis e incompatíveis, da abordagem pragmática com as recomendações presentes nas normas para elaboração de tesauros ANSI/NISO Z39.19:2005, ISO25964 – 1(2011) e ISO25964 – 2 (2013).

4.2 A BORDAGEM PRAGMÁTICA NAS NORMAS PARA ELABORAÇÃO DE TESAUROS

A partir da análise que apresentamos na seção anterior elaboramos um quadro de referência para sistematizar a identificação de como os elementos da abordagem pragmática se encontram nas recomendações das normas para elaboração de tesauros.

Os aspectos que caracterizam a abordagem pragmática na OC são em sua maioria de cunho teórico e nas normas praticamente não se encontram explicações teóricas, devido ao seu caráter mais aplicado e

prático. No entanto, as normas apresentam definições, explicações e exemplos de determinados procedimentos que podem demonstrar o entendimento sobre alguns aspectos e temas.

A seguir descrevemos a interpretação comparativa feita, com os argumentos sobre o que foi considerado compatível e o que foi considerado incompatível. Limitamo-nos a apresentar as recomendações e demais declarações das normas que apresentavam alguma ligação com os enunciados da abordagem pragmática na OC. Ou seja, não descrevemos todos os aspectos das normas, apenas aqueles que subsidiaram a nossa análise comparativa, nos limitando assim a descrever o que se mostrou compatível ou incompatível com a abordagem pragmática.

Para facilitar a apresentação dessa comparação descrevemos os 9 enunciados, elaborados a partir da análise apresentada na seção anterior.

- 1 - A informação é um fenômeno construído;
- 2 - contextos múltiplos e heterogêneos e dilemas éticos;
- 3 - a qualidade da informação não é constante;
- 4 - a prática é importante para a construção do significado;
- 5 - a interação é importante para a validação;
- 6 - a bibliografia como fonte para identificar pontos de vista distintos;
- 7 - comunicar é mais importante do que representar;
- 8 - guiar para uma escolha informada;
- 9 - relações semânticas para contextualizar os pontos de vista distintos.

4.2.1 ANSI/NISO Z39.19:2005

A norma ANSI/NISO Z39.19 (2005) é composta por 11 capítulos: 1) Introdução, 2) Escopo, 3) Normas referenciadas, 4) Definições, abreviações e siglas, 5) Vocabulários controlados – finalidade, conceitos, princípios e estrutura, 6) Termos – escolha, escopo e forma, 7) Termos compostos, 8) Relações, 9) Apresentação de vocabulários controlados, 10) Interoperabilidade e 11) Construção, teste, manutenção e sistemas de gestão. Apresenta, ainda, diversos apêndices, bibliografia, glossário e índice.

Um dos principais objetivos de um tesauro, de acordo com a norma ANSI/NISO Z39.19 (2005, p. 12):

É garantir que cada conceito distinto seja descrito por uma forma linguística única. Se existirem várias formas, essas devem ser controladas ou regularizadas para que a informação ou conteúdo que são fornecidos a um usuário não sejam transmitidos por vários pontos de acesso, mas estejam reunidos em um só lugar. Por meio do controle de vocabulário, elimina-se a ambiguidade e sinonímia e se assegura que cada termo tenha apenas um sentido e que apenas um termo possa ser usado para representar determinado conceito ou entidade.

Destacamos a preocupação com a representação, o que, conforme vimos, é incompatível com a abordagem pragmática, em especial com a maior preocupação que ocorre nessa abordagem em possibilitar uma comunicação efetiva entre o usuário e o SRI (enunciado 7) por meio da apresentação dos diferentes pontos de vista e modos de uso da linguagem (enunciado 6).

Contudo, a norma também esclarece que:

Uma grande organização pode usar os mesmos termos em diferentes contextos. Nesses casos, devem-se tomar cuidados para garantir que os usuários, tanto os indexadores quanto os pesquisadores, possam fazer as distinções necessárias entre os termos. (ANSI/NISO Z39.19, 2005, p. 13).

A forma para fazer o controle das ambiguidades (palavras homógrafas ou polissêmicas), segundo a norma ANSI/NISO Z39.19(2005) seria com a adição de um qualificador, uma palavra adicional colocada entre parênteses logo após o termo, que visa esclarecer o sentido do mesmo. Apresenta-se como exemplo o termo mercúrio, o qual poderia ter como qualificador automóvel, planeta, metal e mitologia. A estrutura do VC, com relações hierárquicas, também é apontada como meio para controlar as ambiguidades.

No entanto, as ambiguidades que ocorrem na bibliografia científica não são tão bem determinadas/definidas e estáveis como a ambiguidade do exemplo apresentado. Nesse sentido encontramos na norma ANSI/NISO Z39.19 (2005, p. 14) a recomendação referente ao

emprego da análise facetada, a qual é considerada particularmente útil para:

- campos novos e emergentes em que não há conhecimento completo sobre o domínio ou as relações entre os objetos de conteúdo são desconhecidas ou não muito bem definidas;
- áreas interdisciplinares, em que há mais de uma perspectiva sobre como olhar determinados assuntos, ou quando são necessárias combinações de conceitos;
- vocabulários em que várias hierarquias são necessárias, mas podem ser insuficientes devido às dificuldades em definir os seus limites claros; ou
- classificação de documentos eletrônicos, em que a localização e alocação em uma estante, por exemplo, não é uma questão importante.

A análise facetada, com base no trabalho de Ranganathan, é considerada útil especialmente em tesauros que possuem milhares de termos. Na análise facetada a abordagem adotada é a *bottom-up*, enquanto que uma organização puramente hierárquica seria orientada por disciplinas, em que as áreas que agrupam semelhantes já são determinadas, e a abordagem é, portanto, a *top-down* (ANSI/NISO Z39.19, 2005).

A preocupação que a norma apresenta para recomendar a análise facetada⁹¹ nos parece semelhante àquela de Hjørland (2007b, 2003) e Hjørland e Pedersen (2005), dentre outros, uma forma de organizar e facilitar o acesso a uma bibliografia que apresente pontos de vistas distintos. Essa recomendação está de acordo com a visão pragmática na OC, especificamente com o enunciado 7.

O método *bottom-up* é recomendado também na abordagem comitê, que é uma das três abordagens recomendadas para a construção de VC na norma. Além da abordagem comitê, a norma recomenda a abordagem empírica e a combinação dos métodos.

Na abordagem comitê:

⁹¹ Ainda que a análise facetada possa ser feita de diferentes maneiras, nos referimos aqui à preocupação apresentada na norma que impulsionou a recomendação da análise facetada.

Especialistas no assunto de domínio do vocabulário controlado elaboraram uma lista dos termos-chave no campo e indicam as relações entre eles, com a ajuda de especialistas em elaboração de vocabulário controlado. As listas de termos podem ser tomadas a partir de várias fontes ou enviadas pelos usuários e comunidades. (ANSI/NISO Z39.19, 2005, p. 91).

Nessa abordagem descreve-se a existência de dois métodos principais, o *top down* e o *bottom up*.

a) Top Down - Os termos mais amplos são identificados em primeiro lugar e, em seguida, termos mais específicos são selecionados para alcançar o nível desejado de especificidade. As estruturas e as relações hierárquicas necessárias são criadas com o prosseguimento do trabalho.

b) Bottom up - Esse caso ocorre com frequência quando listas de termos foram extraídos de um corpus e são posteriormente incorporados no vocabulário controlado. Como no caso acima, as estruturas e as relações hierárquicas necessárias são criadas com o prosseguimento do trabalho, mas a partir dos termos mais específicos para os mais gerais.

Se um novo vocabulário controlado está sendo criado, a abordagem “*top down*” é a preferida, uma vez que em um vocabulário controlado, a abordagem “*bottom-up*” é mais frequentemente usada para adicionar novos termos para cobrir novos conceitos. (ANSI/NISO Z39.19, 2005, p.91).

Na abordagem empírica os dois métodos básicos são descritos como método dedutivo e método indutivo:

a) No método dedutivo - termos são extraídos de documentos (por seres humanos ou computadores), opcionalmente durante a fase preliminar de indexação, mas não é feita nenhuma tentativa de controlar o vocabulário, nem para determinar as relações entre os termos, até que um número suficiente de termos tenha

sido coletado. Todos os termos são revisados por um grupo de profissionais, de preferência constituídos por especialistas da área e especialistas em informação. Primeiramente identificam-se os termos que representam as classes mais amplas, e depois alocam-se os termos restantes nessas classes, com base em suas relações lógicas, de modo que as hierarquias tendem a ser estabelecidas de conceitos mais amplos para os mais específicos. O controle de vocabulário deve ser desenvolvido na fase em que as hierarquias e outras relações são estabelecidas.

b) No método indutivo - novos termos são selecionados para potencial inclusão no vocabulário controlado, da maneira como são encontrados em documentos. O controle de vocabulário é aplicado desde o início. Se o vocabulário que está sendo construído tem algum tipo de arranjos hierárquico, cada novo termo admitido é designado como membro de uma ou mais classes que são construídas numa base *ad hoc* em estágio inicial. O vocabulário controlado é, portanto, determinado do específico para o geral. A construção do vocabulário controlado é considerada desde o início como uma operação contínua. A assistência de especialistas no assunto é altamente recomendável; os peritos atuam como membros de um conselho editorial formal ou comitê. (ANSI/NISO Z39.19, 2005, p.91 e 92).

Com relação à combinação dos métodos, que seria a terceira abordagem, afirma-se que:

Na prática, mais do que uma dessas abordagens podem ser utilizadas em uma fase ou outra, durante a construção de um vocabulário controlado. Por exemplo, hierarquias e outras relações entre os termos que foram inicialmente estabelecidas indutivamente poderiam mais tarde ser examinadas a partir de um ponto de vista dedutivo. Ambas as técnicas são essencialmente empíricas, e, portanto, aceita-se que algumas decisões a respeito dos termos e suas inter-relações que foram feitas durante os primeiros

estágios de compilação podem ser revistas, com a experiência. Os compiladores devem verificar os termos e hierarquias com frequência para assegurar a aplicação coerente dos princípios adotados para as relações entre termos e o desmembramento de termos compostos. (ANSI/NISO Z39.19, 2005, p. 92).

As recomendações para a construção de VC são abrangentes e apresentam diferentes abordagens e métodos. Podemos dizer que a abordagem comitê é incompatível com o que analisamos acerca da abordagem pragmática, enunciados 6 e 9, uma vez que leva em conta principalmente a opinião de especialistas e usuários, não a literatura da área. Na abordagem empírica os termos são identificados na literatura da área e este aspecto está de acordo com a abordagem pragmática. No entanto, nos demais esclarecimentos quanto à forma de análise para a identificação dos elementos do VC, tanto no método indutivo quanto no dedutivo, a classificação parece se dar a partir de generalizações definidas *a priori*, até mesmo no método indutivo. Ou seja, as classes são determinadas antes da análise da literatura, o que é incompatível com a abordagem pragmática.

De acordo com a visão pragmática as categorias ou classes e demais aspectos não são definidos *a priori*, mas sim após a análise da bibliografia da área. Retomando o princípio de garantia literária como forma para se identificar os aspectos de uma determinada área que devem constar em um SOC.

A norma ANSI/NISO Z39.19 (2005, p. 6) apresenta a seguinte definição para garantia literária: “Justificativa para a representação de um conceito em uma linguagem de indexação ou para a seleção de um termo preferido devido a sua ocorrência frequente na literatura.”

Orienta para o uso da remissiva, ver também, nos casos de garantia organizacional e garantia do usuário. A diferença dessas em relação à garantia literária é o motivo da justificativa, devido à características e ao contexto da organização, e devido aos frequentes pedidos de informações sobre o conceito ou pesquisas sobre o termo pelos usuários.

A necessidade de escolha de um termo preferido ocorreria quando “um conceito pode ser representado por vários termos que têm significados iguais ou semelhantes” (ANSI/NISO Z39.19, 2005, p. 13). O termo mais usado entre os usuários, na literatura e no contexto organizacional seria o termo preferido e os demais termos seriam os

termos não preferidos. Nesse caso se deve fazer a ligação entre os termos, por meio de uma relação de equivalência.

Nos três tipos de garantias citadas a preocupação está em justificar a representação de um conceito ou a escolha de um termo preferido devido a maior ocorrência na literatura, no contexto organizacional ou entre os usuários. Ou seja, seria uma espécie de análise estatística da frequência dos termos e dos conceitos, com a qual se escolheria aqueles com maior frequência, o que poderia excluir ainda mais os pontos de vista e modos de expressão das minorias, o que está em desacordo com a visão pragmática na OC, em especial com os enunciados 3 e 6. Afirma-se também que “termos preferenciais devem ser selecionados para atender às necessidades da maioria dos usuário” (ANSI/NISO Z39.19 2005, p.44).

A partir disso podem ocorrer duas situações nos SRI: usar o termo preferido na busca, caso contrário não se recupera nada (nesse caso o usuário também poderia ter acesso ao tesouro ou outro VC e fazer a busca pelo termo desejado verificando se existe algum termo sinônimo considerado preferido); ou, no caso de SRI utilizar um tipo de VC denominado anéis de sinônimos, mesmo a busca pelos termos sinônimos (e não preferidos) recuperaria todos os documentos indexados com o termo preferido. O que poderia gerar uma baixa precisão, e necessitaria de recursos adicionais para a seleção após a busca inicial.

As recomendações que a norma apresenta, de maneira geral, visam melhorar a precisão e a revocação na RI. A revocação seria melhorada com métodos para o controle do vocabulário: termos preferidos e relações de equivalência para controle de sinônimos; forma do termo preferencial (substantivos, etc.); relações associativas; relações hierárquicas e classificação; pós-coordenação; mapeamento de conceitos / agrupamento. A precisão seria melhorada com métodos para o controle do vocabulário: qualificadores para controlar a ambiguidade; relações hierárquicas; termos compostos; pré-coordenação. A norma ANSI/NISO Z39.19(2005) apresenta a seguinte definição para precisão:

Medida da capacidade de um sistema de busca recuperar apenas documentos relevantes. Normalmente expressa como uma percentagem calculada pela divisão do número de documentos relevantes recuperados pelo número total de documentos recuperados. Uma busca de alta precisão assegura que, na sua maior parte, os documentos recuperados serão relevantes. No

entanto, uma pesquisa de alta precisão pode não recuperar todos os documentos relevantes. (ANSI/NISO Z39.19, 2005, p. 8).

E a seguinte definição para revocação:

Medida da capacidade de um sistema de busca recuperar todos os documentos com informações relevantes. Normalmente expressa como uma percentagem calculada pela divisão do número de documentos recuperados pelo número total de documentos relevantes em uma coleção. A pesquisa de alta revocação recupera um conjunto abrangente de documentos relevantes da coleção. No entanto, uma revocação elevada aumenta a possibilidade de que documentos irrelevantes também sejam recuperados. (ANSI/NISO Z39.19, 2005, p. 8).

Apesar de se referir ao conceito de relevância nessas definições de dois conceitos centrais para as suas diversas recomendações, bem como em outras recomendações, a norma ANSI/NISO Z39.19 (2005) não apresenta uma definição do que é relevância. No entanto, na passagem a seguir podemos notar que a ideia de relevância presente na norma diz respeito à visão tradicional e cognitivista de que aquilo que o usuário espera e deseja encontrar é o que é relevante:

Vocabulários controlados servem a cinco propósitos:

- 1. Tradução:** fornecer um meio para converter a linguagem natural dos autores, indexadores e usuários do vocabulário, que podem ser utilizados na indexação e na recuperação da informação.
- 2. Consistência:** promover a uniformidade no formato dos termos e na atribuição de garantias.
- 3. Indicação de relações:** indicar relações semânticas entre os termos.
- 4. Rotular e navegar:** fornecer hierarquias consistentes e claras em um sistema de navegação para ajudar os usuários a localizar os documentos desejados.
- 5. Recuperação:** servir como um auxiliar de pesquisa na localização de objetos de conteúdo. (ANSI/NISO Z39.19, 2005, p. 11).

É interessante notar que, apesar de apresentar uma preocupação em guiar o usuário (navegação), o objetivo ainda diz respeito a mostrar como encontrar aquilo que o usuário espera, o que deseja, não algo que o ajude a alcançar um objetivo específico (que pode ser algo inesperado e novo), o que é incompatível com a abordagem pragmática em especial no que diz respeito ao enunciado 8.

Para nós, esse aspecto evidencia a presença da influência do paradigma cognitivo, assim como na definição de conceito como:

Uma unidade de pensamento, formado pela combinação mental de algumas ou de todas as características de um objeto real ou imaginário, concreto ou abstrato. Conceitos existem na mente como entidades abstratas independentes dos termos usados para expressá-los. (ANSI/NISO Z39.19, 2005, p. 158).

Essa definição é oposta ao entendimento que ocorre na abordagem pragmática de conceitos como significados socialmente negociados e, ainda, à visão de que o conhecimento e a subsequente construção da informação não dependem apenas de fatores subjetivos, mas que a interpretação da realidade está ancorada no contexto social, em que os significados se dão a partir do uso da linguagem, conforme o enunciado 1.

Apesar da definição de conceito que é apresentada ser incompatível com a abordagem pragmática, nas recomendações quanto à abrangência dos conceitos de um VC a norma afirma que:

O escopo dos termos é restrito aos significados selecionados no domínio do vocabulário controlado. Cada termo deve ser formulado de tal forma que transmita o escopo pretendido para qualquer usuário do vocabulário controlado. [...] Termos cujos significados se sobrepõem em uso e homógrafos no geral (termos com grafias idênticas, mas significados diferentes) devem ser evitados tanto quanto possível, na seleção dos termos. No entanto, a garantia literária, ou seja, a ocorrência de termos em documentos e outros critérios podem justificar o uso de tais termos. (ANSI/NISO Z39.19, 2005, p. 20).

O que leva em consideração o uso da linguagem para a definição do escopo dos termos e seus significados e é compatível com a abordagem pragmática, enunciado 4, e aponta certa incoerência entre algumas afirmações da norma.

Para transmitir o escopo pretendido de um termo a norma ANSI/NISO Z39.19 (2005) apresenta como opções a inclusão de qualificadores, notas de escopo e notas históricas. É esclarecido, no entanto, que o uso de qualificadores deve ser evitado sempre que possível devido aos problemas que parênteses podem causar na recuperação, pois um qualificador faz parte do termo. Uma nota escopo, por sua vez, pode ser usada para restringir ou expandir a aplicação de um termo, distinguir entre os termos que têm significados sobrepostos em linguagem natural, ou fornecer conselhos sobre o uso de um termo a qualquer indexador ou pesquisador. Uma nota escopo deve indicar o significado de um termo escolhido e pode também indicar outros significados que são reconhecidos em linguagem natural, mas que tenham sido deliberadamente excluídos do vocabulário controlado. Ao contrário de um qualificador entre parênteses, uma nota de escopo não é parte de um termo. Enquanto qualificadores são geralmente adicionados apenas para homógrafos, uma nota de escopo pode ser fornecida a qualquer termo. Semelhante à nota de escopo a nota histórica é usada para controlar o desenvolvimento de termos ao longo do tempo e oferecem importante orientação para os pesquisadores que estão interessados em um tema que abrange muitas décadas. É especialmente importante para indicar quando e como um termo mudou ao longo do tempo. Se for o caso, uma nota histórica também pode incluir o termo sucessor, e/ou o termo predecessor. (ANSI/NISO Z39.19 2005, p. 22).

As recomendações referentes às notas de escopo e histórica são compatíveis com a abordagem pragmática, enunciados 3 e 2, ao passo que consideram a possibilidade de diferenças conceituais e terminológicas devido a aspectos históricos e contextuais.

Conforme descrevemos na subseção 4.1.2 essas diferenças conceituais acabam gerando questões éticas para a elaboração de um tesouro e demais SOC. Na norma ANSI/NISO Z39.19 (2005) encontramos referência aos aspectos éticos para a seleção dos termos preferenciais que:

Devem refletir o uso de pessoas familiarizadas com o domínio do vocabulário controlado levando em conta a garantia literária, organizacional e do

usuário. Termos neutros devem ser selecionados, por exemplo, as nações em desenvolvimento, em vez de países subdesenvolvidos. (ANSI/NISO Z39.19, 2005, p. 20).

O que nos deixou em dúvida quanto ao que fazer se o domínio do conhecimento não for neutro, ou seja, se a literatura, a organização e os usuários apresentarem algum tipo de viés. O exemplo apresentado é simplista e a decisão a ser tomada ali é clara, mas diferente daquelas que os profissionais responsáveis pela elaboração de um SOC poderão ter que tomar na vida real. Ao longo da norma não localizamos maiores referências quanto aos aspectos éticos.

A norma ANSI/NISO Z39.19 (2005, p. 81) recomenda o uso de “descrições completas do vocabulário controlado e seu uso, que podem compor documentação separada para os mantenedores de vocabulário, indexadores e pesquisadores”. Essa documentação deve trazer esclarecimentos quanto aos diversos aspectos do VC: a finalidade; o escopo, ou seja, o campo de assunto coberto, com áreas centrais e marginais identificadas separadamente; o significado de todas as convenções, abreviaturas, e quaisquer sinais de pontuação utilizados; as regras e as autoridades adotadas na seleção das formas preferidas de termos e no estabelecimento de suas relações; se o vocabulário controlado foi elaborado em conformidade com uma norma nacional ou internacional; as regras de arquivamento empregadas, citando uma norma ou orientação adequada, quando utilizada; o número total de termos, com os totais separados para os termos preferenciais e de entrada; a data em que o vocabulário controlado foi atualizado; uma declaração sobre a política de atualização; a informação de contato (por exemplo, nome e endereço) do responsável pela organização à qual comentários e sugestões podem ser enviados; e as convenções de navegação on-line especiais ou opções especiais. (ANSI/NISO Z39.19 2005).

Assim, consideramos que esta recomendação é compatível com a abordagem pragmática, enunciado 2, ao passo que visa o esclarecimento quanto aos diversos aspectos do VC, com esses esclarecimentos seria possível determinar explicitamente a perspectiva representada no sistema, o que é importante pois “se não podemos eliminar o viés, então devemos nos esforçar para ser mais responsáveis sobre ele” (HJØRLAND, 2008a).

Quanto à atualização dos VC a norma descreve que:

A atualização dos vocabulários controlados é reflexo da linguagem, e eles são, portanto, instrumentos dinâmicos. Políticas e procedimentos devem ser estabelecidos para a revisão periódica da terminologia, o estabelecimento de novos termos, e substituição de termos obsoletos, especialmente em áreas onde a terminologia muda rapidamente (ANSI/NISO Z39.19, 2005, p. 96).

O que é compatível com a abordagem pragmática, enunciado 3, uma vez que considera os VC como instrumentos dinâmicos devem refletir as mudanças da linguagem do domínio.

A atualização se dá, dentre outras formas, por meio da inclusão e eliminação de elementos do VC. Quanto à eliminação de termos a norma ANSI/NISO Z39.19 (2005, p. 97) recomenda que:

Termos que foram atribuídos em demasia e com pouca frequência na indexação devem ser considerados candidatos para alteração ou supressão, pois os dois tipos de termos são geralmente ineficazes na recuperação. Em alguns casos, um termo usado em demasia pode ser substituído por dois ou mais termos de maior especificidade, ou estes podem ser adicionados à hierarquia do termo como termos mais específicos.

Nessa recomendação não fica claro, ou pelo menos não é dito de forma explícita, mas pode ser que o uso demasiado para a indexação de um determinado termo do VC seja visto como negativo devido ao fato de que uma busca com esse termo gera uma alta revocação e possivelmente baixa precisão (pois teoricamente quando uma aumenta a outra diminui). Em alguns casos, termos mais específicos podem realmente resolver o problema, mas se o nível de especificidade já for o adequado e o assunto que aquele termo representa é altamente abordado na literatura, uma nova forma para melhorar a precisão na recuperação teria talvez de ser implementada. Uma alternativa seria aumentar as possibilidades de combinações de termos de busca, e para isso outros tipos de relações, além das hierárquicas, poderiam contribuir.

Esses outros tipos de relações poderiam contribuir ao permitir novos reagrupamentos a partir de uma busca inicial, o que possibilita

refinamentos dos resultados da busca e possibilita maior precisão a partir das escolhas dos usuários, de modo mais fácil e intuitivo. Na norma ANSI/NISO Z39.19 (2005) este tema é tratado no capítulo sobre interoperabilidade⁹², mais especificamente sobre a interoperabilidade de relações semânticas.

Uma vez que as diversas relações entre os termos em vários vocabulários controlados são identificadas, algo deve ser feito para reter e manter essas relações para uso futuro. Várias opções estão sendo usadas para armazenar essas informações e disponibilizá-las em apoio a iniciativas de interoperabilidade. (ANSI/NISO Z39.19, 2005, p. 86, tradução nossa).

Dentre essas destacam-se as redes semânticas e os bancos de dados lexicais.

Uma rede semântica pode ser utilizada para agrupar termos de um ou diversos vocabulários controlados de acordo com uma estrutura organizacional de base que é comum a todos os vocabulários controlados candidatos à adesão. Isso normalmente é alcançado pelo agrupamento de termos de acordo com a sua relação com algum esquema conceitual. (ANSI/NISO Z39.19, 2005, p. 87).

⁹² Dois interesses conflitantes impulsionam o desenvolvimento e uso de vocabulários controlados. Por um lado, as comunidades desenvolvem vocabulários controlados específicos para seus conceitos, terminologia e necessidades, como meio de acesso aos dados que são importantes e para codificar o conteúdo específico para o seu domínio. Por outro lado, pesquisadores e aqueles que os apoiam querem usar uma única pesquisa para encontrar recursos em bancos de dados que servem a diferentes domínios e são acessados por diferentes vocabulários controlados, por meio dos quais não se admite qualquer consenso sobre conceitos, terminologia e organização do conhecimento. A interoperabilidade é o objetivo de levar vocabulários que na maioria dos casos foram destinados a uso isolado e relacioná-los uns com os outros para permitir que pesquisas extraídas de um vocabulário possam ser eficazes no outro; recursos descritos com um vocabulário possam ser descritos com sucesso em outro; e vocabulários vinculados a domínios específicos possam ser reunidos em mapas conceituais e terminológicos maiores e mais abrangentes. (ANSI/NISO Z39.19, 2005, p.82, tradução nossa)

Ao estabelecer uma rede semântica, é fundamental que os tipos de conceitos e tipos de relacionamentos sejam definidos. No *WordNet*, por exemplo, identificam-se 135 tipos semânticos e 54 relacionamentos. Por vezes, a rede semântica é derivada a partir dos próprios vocabulários controlados e noutros casos é derivada a partir dos bancos de dados de destino. Nos casos em que uma rede semântica é derivada do banco de dados subjacente, um resultado pode ser a geração de anéis de sinônimos para o uso durante a recuperação. Quando esses estão disponíveis, os pesquisadores normalmente têm a opção de seleccionar a semântica de proximidade dos termos desejada a ser utilizada na pesquisa. As escolhas, por exemplo, podem ser: sinônimos exatos; quase sinônimos; termos estreitamente relacionados; termos vagamente relacionados; antônimos. (ANSI/NISO Z39.19, 2005, p.87, tradução nossa).

Um banco de dados lexical, por sua vez, “pode ser usado para associar termos de vários vocabulários controlados em *clusters* de conceitos relacionados. Vários tipos de relações podem ser acomodadas incluindo sinônimos, antônimos e relações hierárquicas de vários tipos” (ANSI/NISO Z39.19, 2005, p. 89). O *WordNet* é apresentado também como um exemplo de banco de dados lexical no qual uma busca pelo termo *bridge* (ponte) apresenta os diversos significados possíveis, e a partir da escolha de um dos significados o usuário pode então ver os sinônimos do termo, coordenar termos, ver as relações, o domínio, a familiaridade, etc. Outro exemplo de banco de dados lexical é os *Xrefer* que usa uma abordagem gráfica para exibir os relacionamentos entre conceitos.

Consideramos que essas recomendações quanto ao armazenamento e manutenção de relações para a interoperabilidade são compatíveis com a abordagem pragmática, especialmente com o enunciado 8, uma vez que possibilitam uma visão geral e as possíveis opções de escolha para o usuário. Ainda que a forma para identificar os diferentes tipos de relações não seja descrita e o exemplo que foi apresentado, o *WordNet*, ser considerado por Hjørland (2013b) como um exemplo da abordagem cognitiva e com base em pesquisa psicológica e em Hjørland (2007b) afirmar-se que na literatura disponível é difícil determinar as bases sobre as quais as decisões no

WordNet são feitas. O que nos faz entender que as relações que poderiam ajudar a guiar o usuário na sua escolha podem não ser compatíveis com a abordagem pragmática, conforme o enunciado 9, mas a ideia geral de guiar o usuário é compatível.

4.2.2 ISO25964 – 1

A norma ISO25964 *Thesauri and interoperability with other vocabularies* é dividida em duas partes. A parte 1: *Thesauri for information retrieval*, foi publicada em 2011 e apresenta recomendações para o desenvolvimento e manutenção de tesouros, monolíngues ou multilíngues.

Essa primeira parte da norma é composta pelos seguintes capítulos: Introdução, 1) Escopo, 2) Termos e definições, 3) Símbolos, abreviações e outras convenções, 4) Tesouro - visão geral e objetivos, 5) Conceitos e seu escopo em um tesouro, 6) Termos do tesouro, 7) Conceitos complexos, 8) A relação de equivalência em um contexto monolíngue, 9) Equivalência entre linguagens, 10) Relações entre conceitos, 11) Análise facetada, 12)) Apresentação e *layout*, 13) Gestão da construção e manutenção do tesouro, 14) Guia para *software* de gerenciamento de tesouro, 15) Modelo de dados, 16) Integração do tesouro com aplicativos, 17) Formatos de intercâmbio, 18) Protocolos.

A norma ISO25964-1 define que o objetivo global de um tesouro é:

Orientar o indexador e o usuário a escolher o mesmo termo para o mesmo conceito. Os conceitos são representados por termos, e para cada conceito, uma das possíveis representações é selecionada como o termo preferido. Em segundo lugar, um tesouro deve apresentar os termos preferidos de forma que as pessoas identifiquem facilmente o que necessitam. Isto é alcançado pelo estabelecimento de relações entre os termos – e/ou entre os conceitos – usando as relações para apresentar os termos de forma estruturada. (ISO25964-1, 2011, p.15, tradução nossa)

Além disso, é esclarecido que em sistemas de pesquisas ou sistemas de indexação os termos e as relações dos tesouros podem ser explorados de inúmeras maneiras, tais como:

Expansão da pesquisa, sugestão de termos de pesquisa alternativos, suporte para agrupamento ou outros meios para refinar a pesquisa, identificação de erros de ortografia comuns, suporte para indexação automática\”. (ISO25964-1, 2011, p.15, tradução nossa).

O objetivo dos tesauros e as maneiras adicionais que seus termos e relações podem ser utilizados estão de acordo com a abordagem pragmática, no que diz respeito ao enunciado 7, pois existe uma preocupação em possibilitar a comunicação, em proporcionar a compreensão por parte do usuário da linguagem utilizada no SRI. Com a efetiva comunicação, por meio do uso da mesma linguagem, descreve-se que os tesauros também poderiam contribuir com a expansão ou com o refinamento da pesquisa, guiando o usuário à novas possibilidades, o que também é compatível com a abordagem pragmática conforme o enunciado 8. Para este processo a importância atribuída às relações semânticas, que constam como meio para alcançar o objetivo do tesouro de apresentar um *display* estruturado, também é adequada e compatível com a abordagem pragmática conforme enunciado 9.

O controle do vocabulário é considerado necessário para alcançar o objetivo global, de orientar o indexador e o usuário a escolher o mesmo termo para o mesmo conceito. Este controle, segundo a norma ISO25964-1 (2011), se faz da seguinte maneira:

- a) Conceitos e termos são deliberadamente restritos a um escopo para a seleção de significados. [...] em um tesouro é geralmente restrito um único significado que serve para as necessidades de maior efetividade de um sistema de recuperação. A estrutura de um tesouro, seu *display* de relações hierárquicas, frequentemente indica o significado pretendido para um termo. Se esta técnica não for suficientemente explícita, uma nota de escopo deve acompanhar o termo. Esta nota deve mostrar o significado escolhido, e pode também indicar outros significados que são reconhecidos na linguagem natural mas que foram deliberadamente excluídos para os propósitos da indexação.
- b) Quando o mesmo conceito pode ser expresso por dois ou mais sinônimos ou quase-sinônimos na mesma linguagem, um destes termos é

normalmente selecionado como o termo preferido, o qual é então usado consistentemente na indexação [...] Uma referência para o termo preferido deve ser feita a partir de cada sinônimo que também podem funcionar como pontos de acesso para o usuário.

Uma consequência de utilizar as medidas a) e b) para o controle do vocabulário é que a linguagem resultante pode não corresponder às preferências dos usuários. O tesauro tem um importante papel na mediação entre os termos usados no discurso e aqueles que funcionam efetivamente para a recuperação da informação. Para alcançar os benefícios com a recuperação, os usuários precisam aceitar certo grau de artificialidade do vocabulário (embora em alguns sistemas, esta dificuldade pode ser superada pela substituição automática do termo dos usuários com a forma preferida).

[...] Se o tesauro é simétrico como descrito nesta parte da ISO25964 e modelado conforme capítulo 15, cada conceito deve ter um termo preferido em cada linguagem, e o escopo deve ser o mesmo em todas as linguagens. Esta limitação às vezes adiciona artificialidade ao vocabulário. Abordagens alternativas (por exemplo o uso de mapeamentos entre vocabulários) serão descritas na ISO 25964-2. (ISO25964-1, 2011, p. 16, tradução nossa).

Efetuar o controle do vocabulário é considerado necessário para orientar o indexador e o usuário a escolher o mesmo termo para o mesmo conceito para fins de recuperação da informação. Esse controle diz respeito a uma adaptação da linguagem natural, que possui aspectos que são tradicionalmente considerados prejudiciais à efetividade da recuperação da informação. Nesse sentido a ISO25964-1 é incompatível com a abordagem pragmática, conforme o enunciado 7, pois apesar de no objetivo se descrever uma preocupação com a comunicação, com o entendimento por parte do usuário da linguagem de indexação usada no SRI, a forma de se alcançar isso diz respeito ao que é considerado como essencialismo na abordagem pragmática.

Para a construção de um tesauro diversas etapas são descritas como necessárias, a primeira diz respeito ao planejamento e definição dos objetivos, que devem levar em conta os seguintes aspectos:

- a) como o tesauro será utilizado, e por quem;
 - b) se existem restrições impostas pelas limitações do *software* com o qual será usado;
 - c) o nível de conhecimento dos usuários na área temática do tesauro e no uso de tesouros.
- Diferentes partes interessadas (tais como executivos de gestão, a equipe de apoio em TI e os futuros usuários) têm expectativas diferentes sobre o que um tesauro é e como ele vai funcionar. Uma definição dos objetivos devem ser acordados por todas as partes como passo fundamental para alcançar um trabalho em equipe eficaz.(ISO25964-1, 2011, p. 88, tradução nossa).

O último aspecto, sobre levar em conta o nível de conhecimento dos usuários como item central na definição dos objetivos do tesauro, nos chama a atenção, pois é passível de comparação com a abordagem pragmática na OC. Esse aspecto, por sua vez, é criticado na abordagem pragmática, especialmente por Blair (2003), Hjørland e Christensen (2003) e Frohmann (1990), ao se considerar que isso pode gerar limitações. Como o tesauro poderá ampliar a visão do usuário sobre determinado tema se for baseado naquilo que o usuário já sabe? Ou, ainda, naquilo que ele deseja encontrar, que é o entendimento tradicional para relevância, conceito nos quais os índices de precisão e revocação estão ancorados. O enunciado 8 foi proposto com base nessa discussão, mas, orientar para uma escolha informada parecer ser incompatível com a recomendação descrita. O nível de conhecimento dos usuários poderia ser um indicador para definir a política de desenvolvimento de coleções. Coleção que seria analisada para o desenvolvimento do tesauro.

A pouca importância atribuída às coleções (à bibliografia que será indexada com o tesauro) para a construção de um tesauro é evidenciada também na descrição dos recursos que devem ser levados em conta na fase de planejamento e nos estágios iniciais de compilação. Os recursos são divididos em humanos e de vocabulário, quanto aos recursos de vocabulário descreve-se:

- a) Existência de tesouros ou esquemas de classificação com um escopo que corresponda parcialmente ou totalmente às exigências. [...].
- b) Coleções de terminologias ou consultas frequentes que os colegas possam ter reunido no decurso do seu trabalho.
- c) Os índices de bases de dados existentes ou outras ferramentas de referência relevantes, mostrando os pontos de entrada úteis para o assunto. [...].
- d) O registro de transações de um site relevante, mostrando os termos mais utilizados pelos usuários.
- e) Obras de referência padrão, assim como dicionários, terminologias, nomeclaturas e enciclopédias são úteis para verificação do significado de um termo e validação de estruturas hierárquicas, mas não como guia para a seleção de termos. (ISO25964-1, 2011, p.89 e 90, tradução nossa).

Quanto aos estágios iniciais de compilação, para a coleta de termos e conceitos, afirma-se que as “principais fontes de termos estão listados de a) a d) no item 13.1.3.3”, o que apresentamos acima. Completa-se a explicação afirmando que “o texto completo da literatura fundamental do escopo exigido também é útil” (ISO25964-1, 2011, p.90, tradução nossa). Nesse sentido também é descrito que técnicas automáticas podem ser usadas, para extrair palavras-chaves e frases a partir da literatura. E considera-se que é relativamente fácil estudar o material a ser introduzido. Assim, a ênfase dada a esse aspecto deveria ser a mesma a quaisquer indicações do que é que os usuários querem buscar, como os registros de consultas. Os usuários também podem ser convidados a sugerir termos e conceitos úteis. (ISO25964-1, 2011).

Conforme descrevemos, a bibliografia para a qual o tesouro será usado na indexação não é considerada um dos recursos principais para a compilação de termos e conceitos, e a sua importância é tida como igual à opinião e consultas anteriores dos usuários, o que é incompatível com a abordagem pragmática, mais especificamente com o enunciado 6.

Além de orientar o indexador e o usuário a escolher o mesmo termo para o mesmo conceito o outro objetivo de um tesouro é apresentar os termos preferidos por meio de relações semânticas de forma que as pessoas identifiquem facilmente o que necessitam. As

relações semânticas consideradas úteis para isso em um tesouro, segundo a norma ISO25964-1(2011, p. 17, tradução nossa), seriam as “relações que são válidas em praticamente todos os contextos, especialmente quando são inerentes aos conceitos que os termos representam, conhecidas como relações paradigmáticas”. Em um tesouro podem ser apresentadas como relações hierárquicas, relações de equivalência e relações associativas, bem como suas possíveis subdivisões. A inclusão de relações sintagmáticas que, segundo a ISO25964-1 (2011, p.16, tradução nossa) são aquelas que “existem apenas por que os conceitos ocorrem juntos no contexto de um documento particular”, não é recomendada.

Conforme descrevemos no referencial teórico dessa pesquisa, de acordo com Khoo e Na (2006), a distinção entre relações paradigmáticas e sintagmáticas é imprecisa e as relações paradigmáticas muitas vezes são expressas no nível sintático. Desse modo, entendemos que a forma como a distinção entre relações paradigmáticas e sintagmáticas foi apresentada, bem como a definição dos tipos de relações úteis para um tesouro (válidas em praticamente todos os contextos e inerentes aos conceitos que os termos representam) indicaria que essas relações poderiam ser estabelecidas por premissas universalistas. Esse aspecto é incompatível com a abordagem pragmática na OC, em especial com o enunciado 9 que, no que se refere às relações semânticas, complementa o enunciado 6. Em ambos se aponta a importância da análise da bibliografia da área para identificação dos elementos de um SOC.

A diferenciação dos tipos de relações semânticas e a definição de quais são úteis em um tesouro, assim como a definição de conceito apresentada, evidenciam certa influência cognitivista e positivista, que são abordagens consideradas opostas à abordagem pragmática na CI e OC. A definição de conceito é semelhante àquela apresentada na ANSI/NISO Z39.19(2005), “unidade de pensamento [...] conceitos podem ser expressos de diversas maneiras. Eles existem na mente como entidades abstratas independentes dos termos usados para expressá-los” (ISO25964-1, 2011, p.3, tradução nossa).

Assim como descrevemos na comparação com as recomendações da ANSI/NISO Z39.19(2005), consideramos que essa definição de conceito é oposta ao entendimento que ocorre na abordagem pragmática de conceitos como significados socialmente negociados e, ainda, à visão de que o conhecimento e a subsequente construção da informação não dependem apenas de fatores subjetivos, mas que a interpretação da

realidade está ancorada no contexto social, em que os significados se dão a partir do uso da linguagem, conforme o enunciado 1.

Outras semelhanças com a norma americana que destacamos, pois foram analisados de forma comparativa com a abordagem pragmática, são: o critério apontado para escolher o termo preferido, “deve ser selecionado para servir às necessidades da maioria dos usuários” (ISO25964-1, 2011, p.47, tradução nossa); a recomendação quanto providenciar uma introdução(ou documentação) com os propósitos; o escopo e demais características do tesouro (ISO25964-1, 2011, p.93); as recomendações referentes às notas de escopo e históricas (ISO25964-1, 2011, p.20 e 23); e o fato de não apresentar uma definição de relevância. Conceito central ao qual é feita referência em diversas recomendações, tanto na parte 1 quanto na parte 2. A seguir prosseguimos com a comparação entre essa segunda parte da ISO 25964 e a abordagem pragmática na OC.

4.2.3 ISO25964 – 2

A segunda parte da ISO25964, *Interoperability with other vocabularies*, foi publicada em 2013 e vem complementar a primeira parte. É aplicável aos tesouros e outros tipos de vocabulários, como taxonomias, esquemas de classificação, listas de cabeçalhos de assunto e anéis de sinônimos que são comumente usados para a recuperação da informação. Descreve, compara e contrasta os elementos e características desses vocabulários que possuem implicações quando a interoperabilidade é necessária. Apresenta recomendações para a criação e manutenção de mapeamentos entre vários tesouros, ou entre tesouros e outros tipos de vocabulários.

É composta por 24 capítulos além da introdução: 1)Escopo, 2) Referências normativas, 3) Termos e definições, 4) Símbolos, convenções e outras abreviações, 5) Objetivos e identificação, 6) Modelos estruturais para o mapeamento entre vocabulários, 7) Tipos de mapeamento, 8) Mapeamentos de equivalência, 9) Mapeamentos de hierarquias, 10) Mapeamentos associativos, 11) Equivalência exata, inexata e parcial, 12) uso de mapeamentos na recuperação da informação, 13) Manuseio de pre-coordenação, 14) Técnicas para identificar mapeamentos candidatos, 15) Gerenciamentos de dados, 16) Apresentação de vocabulários mapeados, 17) Esquemas de classificação, 18) Esquemas de classificação para gerenciamento de registros, 19) Taxonomias, 20) Esquemas de cabeçalho de assunto, 21)

Ontologias, 22) Terminologias, 23) Listas de autoridades, 24) Anéis de sinônimos.

O principal objetivo da interoperabilidade entre vocabulários para fins de RI, segundo a ISO2594-2(2013, p.16, tradução nossa):

É permitir que uma expressão formulada usando um vocabulário seja convertida (ou suplementada) para uma expressão correspondente em outro vocabulário ou mais vocabulários. A expressão em questão pode ser qualquer critério de busca ou parte dos metadados associados a um documento. Em ambos os casos, o mapeamento é o passo fundamental. Se cada conceito no Vocabulário A for mapeado para o conceito(s) correspondente no Vocabulário B, trona-se possível o intercambio (ou ampliação) dos termos ou identificadores representando o conceito em cada um dos vocabulários.

O mapeamento em si, por sua vez, tem como objetivo “mostrar relações entre conceitos em diferentes vocabulários” (ISO25964-2, 2013, p.16, tradução nossa).

Recomendações para os diversos aspectos envolvidos no processo de mapeamento são apresentadas, com diversos exemplos e possibilidades que visam auxiliar nas escolhas necessárias ao se iniciar um projeto de mapeamento. Dentre essas destacam-se a decisão quanto ao modelo geral ou combinação de modelos que serão usados, quanto aos diferentes mapeamentos (hierárquico, associativo, de equivalência e o grau da equivalência), se o mapeamento será utilizado na indexação ou no momento da pesquisa, em qual direção o mapeamento será feito ou se o mapeamento nos dois sentidos é necessário, e como será a mediação humana no processo de conversão.

Diversos aspectos que podem envolver essas decisões são abordados, no entanto, o que nos chama a atenção é o fato dos aspectos descritos na norma se referirem quase que exclusivamente aos vocabulários candidados ao mapeamento ou às características do SRI no qual o mapeamento será aplicado. Como ocorre, por exemplo, na recomendação referente à decisão quanto ao tipo de mapeamento (associativo ou de equivalência inexato) será feito,

A linha que divide um mapeamento associativo e um mapeamento de equivalência inexato é mal-definida e subjetiva, mas pode ser feita pragmaticamente de acordo com o contexto em que os mapeamentos são usados (levando-se em conta, por exemplo, o interesse dos usuários, o volume de recursos relevantes a serem pesquisados, e a capacidade de recuperação e apresentação de distinções feitas no sistema). (ISO25964-2, 2013, p.25, tradução nossa).

Ao longo da norma se faz referência majoritariamente aos vocabulários ou ao processo de RI, assim como na passagem acima apresentada como exemplo, na qual se fala apenas no volume dos recursos relevantes a serem pesquisados.

Entendemos, no entanto, que de acordo abordagem pragmática na OC, apenas com a análise dos recursos a serem pesquisados, ou seja a análise dos documentos a serem pesquisados ou indexados com o mapeamento é que essas e outras decisões poderiam ser tomadas. A capacidade do sistema distinguir as diferenciações, que aparece como pré-determinada na passagem, só seria definida a partir dessa análise. Assim como na passagem apresentada como exemplo, em outras ao longo da norma se passa a impressão que aspectos técnicos podem ser previamente definidos. A partir desses aspectos técnicos ou das características dos vocabulários de origem as demais decisões serão tomadas, o que pode ser problemático, pois ao se mesclar ou unir diferentes contextos em um só, as características individuais consideradas importantes inicialmente para o desenvolvimento do SOC individual provavelmente sofrerão alteração. Analisar apenas os vocabulários de origem pode ser insuficiente.

A ISO25964-2(2013) apresenta importante contribuição para a OC e RI, dada a relevância da questão da interoperabilidade atualmente, mas suas recomendações de forma geral, por dar pouca atenção à bibliografia são incompatíveis com a abordagem pragmática, em especial com o enunciado 6.

No entanto, a importância atribuída às relações semânticas nos mapeamentos, que são considerados o primeiro passo para a interoperabilidade, é notória e vai ao encontro da abordagem pragmática na OC. Na seção 4.3 apresentamos orientações para o estabelecimento de relações semânticas nas quais buscamos unir as recomendações da ISO25964-2, bem como das demais normas, e os princípios da abordagem pragmática na OC. No item a seguir apresentamos a síntese

dos resultados descritos ao longo dessa seção, nos itens 4.2.1, 4.2.2 e 4.2.3.

4.2.4 Síntese da abordagem pragmática nas normas

Conforme descrevemos ao longo da seção 4.2 as normas para elaboração de tesouros, ANSI/NISO Z39.19 (2005), ISO2594-1 (2011) e ISO2594-2 (2013), apresentam tanto recomendações compatíveis quanto incompatíveis com a abordagem pragmática na OC. No quadro a seguir sintetizamos a identificação de como os elementos da abordagem pragmática se encontram nas recomendações das normas para elaboração de tesouros.

Quadro 2 - Síntese de como a abordagem pragmática se encontra nas normas

ABORDAGEM PRAGMÁTICA	RECOMENDAÇÕES DAS NORMAS COMPATÍVEIS	RECOMENDAÇÕES DAS NORMAS INCOMPATÍVEIS
1) A informação é construída, não apenas inventada ou descoberta, e depende do contexto social.		ANSI/NISO Z39.19 (2005, p. 158) e ISO25964-1 (2011, p.3) conceito é uma unidade de pensamento;
2) Os contextos são múltiplos e heterogêneos, o que acarreta dilemas éticos.	(ANSI/NISO, Z39.19, 2005, p. 22) e (ISO25964-1, 2011, p.20 e 23), notas de escopo e históricas; ANSI/NISO Z39.19 (2005, p.81) e (ISO25964-1, 2011, p.93) A documentação/introdução visa o esclarecimento quanto a diversos aspectos do VC, com estes esclarecimentos seria possível determinar explicitamente a perspectiva representada no sistema.	
3) A qualidade da informação não é constante (varia no tempo, um mesmo indivíduo muda sua percepção, e no espaço, entre diferentes indivíduos).	(ANSI/NISO Z39.19, 2005, p. 22) e (ISO25964-1, 2011, p.20 e 23) notas de escopo e históricas; (ANSI/NISO Z39.19, 2005, p.96) considera os VC como instrumentos dinâmicos uma vez que os mesmos devem refletir as mudanças da linguagem do domínio.	
4) A prática é importante para a construção dos significados.	(ANSI/NISO Z39.19, 2005, p.20) leva em consideração o uso da linguagem para a definição do escopo dos termos e seus significados e é compatível com a abordagem pragmática na OC.	

ABORDAGEM PRAGMÁTICA	RECOMENDAÇÕES DAS NORMAS COMPATÍVEIS	RECOMENDAÇÕES DAS NORMAS INCOMPATÍVEIS
5) A interação é importante para validação.		
6) A bibliografia como fonte para identificar pontos de vista distintos.		<p>(ANSI/NISO Z39.19,2005, p. 13; 6) escolha por meio da frequência; ANSI/NISO Z39.19(2005, p. 12) cada termo tenha apenas um sentido e que apenas um termo pode ser usado para representar um determinado conceito ou entidade; (ANSI/NISO Z39.19, 2005, p. 91) Podemos dizer que a abordagem comitê é incompatível uma vez que leva em conta apenas a opinião de especialistas e/ou usuários, não analisa a literatura da área; (ANSI/NISO, 2005, p.91 e 92) as classes são determinadas antes da análise da literatura, o que é incompatível com a abordagem pragmática; A ISO25964-2(2013), de forma geral, dá pouca atenção à bibliografia.</p> <p>(ISO25964-1, 2011, p.90) a bibliografia que o tesauro será usado para indexar não é considerada um dos recursos principais para a compilação de termos e conceitos, e a sua importância é considerada igual à opinião e consultas anteriores dos usuários.</p>

ABORDAGEM PRAGMÁTICA	RECOMENDAÇÕES DAS NORMAS COMPATÍVEIS	RECOMENDAÇÕES DAS NORMAS INCOMPATÍVEIS
7) Comunicar é mais importante do que representar-antiessencialismo.	ANSI/NISO (2005, p.14) facilitar o acesso à uma bibliografia que apresente pontos de vistas distintos e abordagem <i>bottom-up</i> ; (ISO25964-1, 2011, p.15) preocupação em possibilitar a comunicação, em proporcionar a compreensão por parte do usuário da linguagem utilizada no SRI.	ANSI/NISO (2005, p. 12) cada termo tem apenas um sentido e que apenas um termo pode ser usado para representar um determinado conceito ou entidade; (ISO25964-1, 2011, p. 16) apesar de no objetivo se descrever uma preocupação com a comunicação, o entendimento por parte do usuário da linguagem de indexação usada no SRI, a forma para alcançar este objetivo ainda diz respeito ao que é considerado como essencialismo na abordagem pragmática.
8) Guiar para uma escolha informada, e levar em conta que relevante é o que contribui para atingir um objetivo (querer <i>versus</i> precisar).	(ANSI/NISO, 2005, p. 87 e 89) recomendações quanto ao armazenamento e manutenção de relações para a interoperabilidade são compatíveis com a abordagem pragmática. Ao passo que possibilitam uma visão geral e as possíveis opções de escolha para o usuário; (ISO25964-1, 2011, p.15) Com a efetiva comunicação, por meio do uso da mesma linguagem, o SRI também poderia contribuir com a expansão ou com o refinamento da pesquisa, guiando o usuário à novas possibilidades;	(ANSI/NISO – 2005, p. 11) apesar de apresentar uma preocupação em guiar o usuário (navegação), esse objetivo ainda diz respeito a mostrar para ele como encontrar aquilo que ele espera, o que ele deseja; (ISO25964-1, 2011, p.88) Sobre levar em conta o nível de conhecimento dos usuários como item central na definição dos objetivos do tesouro. O que é criticado na abordagem pragmática, especialmente por Blair (2003), Hjørland e Christensen (2003) e Frohmann (1990), pois considera-se que isso pode gerar limitações. Como o tesouro poderá ampliar a visão do usuário sobre determinado tema se for baseado naquilo que o usuário já sabe? Ou, ainda, naquilo que ele deseja encontrar, que é o entendimento tradicional para relevância, conceito nos quais os índices de precisão e revocação estão ancorados.

ABORDAGEM PRAGMÁTICA	RECOMENDAÇÕES DAS NORMAS COMPATÍVEIS	RECOMENDAÇÕES DAS NORMAS INCOMPATÍVEIS
9) Relações semânticas identificadas a partir de estudos epistemológicos para ajudar a guiar o usuário.	(ISO25964-1, 2011, p.15) Com a efetiva comunicação, por meio do uso da mesma linguagem, o SRI também poderia contribuir com a expansão ou com o refinamento da pesquisa, guiando o usuário à novas possibilidades. Para este processo é atribuída uma grande importância às relações semânticas;	<p>ANSI/NISO Z39.19 (2005, p. 87 e 89) As relações que poderiam ajudar a guiar o usuário na sua escolha podem não ser compatíveis com a abordagem pragmática, uma vez que são estabelecidas a partir de estudos psicológicos, como na abordagem cognitiva. (ANSI/NISO, Z39.19 2005, p. 91) Podemos dizer que a abordagem comitê é incompatível uma vez que leva em conta apenas a opinião de especialistas e/ou usuários, não analisa a literatura da área;</p> <p>ISO25964-1(2011, p. 17) a definição dos tipos de relações úteis para um tesouro (válidas em praticamente todos os contextos e inerentes aos conceitos que os termos representam) indicaria que as relações semânticas poderiam ser estabelecidas por premissas universalistas - o que é incompatível com a abordagem pragmática na OC.</p>

Fonte: Elaborado pela autora (2014).

4.3 ORIENTAÇÕES PARA O ESTABELECIMENTO DE RELAÇÕES SEMÂNTICAS

As normas ANSI/NISO Z39.19 (2005) e ISO2594-1 (2011) apresentam recomendações para estabelecer relações semânticas em um VC. Os três tipos principais de relações semânticas, de equivalência, hierárquica e associativa, coincidem em ambas as normas. No quadro a seguir descrevem-se algumas possíveis variações dos diferentes tipos de relações e seus exemplos.

Quadro 3 - Tipos de relações semânticas em tesauros

Tipo de relação	Exemplo
Equivalência	
Sinônimos	ONU / Organização das Nações Unidas
Variantes lexicais	estoque / stock
Quase sinônimos	água do mar / água salgada
Hierárquica	
Genérica	aves / papagaios
Instância	mar/ mar Mediterrâneo
Todo / Parte	encéfalo / tronco encefálico
Associativa	
Causa / Efeito	acidente / lesão
Processo / Agente	medir temperatura / termômetro
Processo / Contra-agente	fogo / retardante de chama

Fonte: ANSI/NISO Z39.19 (2005, p. 42, tradução nossa)

Nas seções a seguir apresentamos, com base nas recomendações das normas, cada tipo de relação de forma mais detalhada e as possibilidades de mapeamento dessas relações para a interoperabilidade entre diferentes vocabulários juntamente com contribuições da abordagem pragmática.

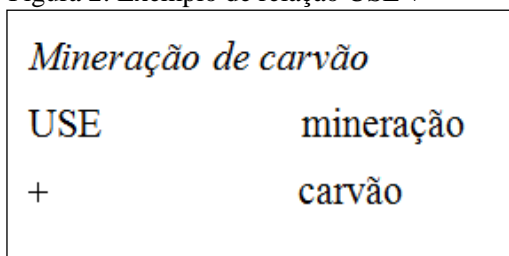
4.3.1 Relações de equivalência

Além dos três tipos de relações de equivalência apresentadas no quadro, sinônimos variantes lexicais e quase sinônimos, a norma

ANSI/NISO Z39.19 (2005) descreve mais dois: postagem genérica, que é recomendada apenas para os termos de áreas periféricas à do VC, ou quando o número de documentos a ser indexado com uma determinada subclasse for muito pequeno, a ponto de não justificar a sua adição ao VC; e referência cruzada para os elementos de um termo composto.

A norma ISO25964-1(2011) apresenta duas subdivisões para as relações de equivalência que coincidem com as da ANSI/NISO Z39.19 (2005), sinônimos e quase-sinônimos. A relação de equivalência que na ANSI/NISO Z39.19 (2005) é denominada postagem genérica e na ISO25964-1(2011, p.48) se fala em “termos específicos agrupados em um conceito mais amplo”. Descreve também um tipo de relação de equivalência que combina dois descritores para representar um conceito complexo e considera o termo composto como não-descritor. Conforme o exemplo a seguir.

Figura 2: Exemplo de relação USE +



Fonte: ISO2594-1(2011)

Além disso, por se tratar de uma norma destinada também à tesouros multilíngues a ISO25964(2011) apresenta recomendações para o estabelecimento de relações de equivalência entre linguagens. Recomenda-se que:

Em um tesouro multilíngue, todas as linguagens tem *status* igual e, se possível, cada conceito deve ser representado em todas as linguagens do tesouro. Uma coleção que tem sido indexada usando uma das linguagens pode ser pesquisada de forma igualmente efetiva usando uma das outras linguagens. (ISO25964-1, 2011, p.50, tradução nossa)

Também se descreve que é possível tratar diferentes dialetos ou sublinguagens como sendo linguagens separadas. Por exemplo, inglês americano, inglês britânico e inglês indiano. De modo semelhante, a terminologia preferida pelos cientistas pode ser apresentada como uma sub-linguagem diferente (ISO25964-1, 2011).

A partir das recomendações das normas e da análise da abordagem pragmática, podemos dizer que as relações de equivalência dependem da escolha do termo preferido e podem contribuir significativamente com a revocação. De maneira geral, os sistemas de busca estão preparados para efetuar a pesquisa apenas com o termo preferido, ou seja, ao buscar por um sinônimo o usuário é orientado a realizar a busca com o termo preferido. Dessa maneira não há distinção entre o termo não preferido e o termo preferido na busca, no entanto, em determinados contextos o significado dos dois termos pode não ser o mesmo e o usuário pode se interessar exatamente por aquilo que diferencia um termo do outro, naquele contexto de uso.

O possível interesse em distinguir os termos que foram considerados equivalentes fica evidente quando pensamos em quase sinônimos e postagem genérica, por exemplo, mas é importante lembrar que sinônimos perfeitos são praticamente inexistentes e que existem implicações éticas em se escolher por uma forma de expressão e não por outra. Consideramos que possibilitar ao usuário distinguir entre o que, para os fins da revocação na RI, foi considerado equivalente não chega a eliminar o viés, mas é um esforço necessário para sermos mais responsáveis sobre ele e determinar explicitamente a perspectiva representada no sistema de informação, conforme Hjørland (2008a) recomenda.

Isso é importante tanto pela dificuldade em se distinguir todos os possíveis contextos de uso já existentes, como também pela dificuldade em se prever os contextos que poderão ocorrer. O significado de um termo não está apenas no presente e no passado, está também no futuro, ou seja, pode mudar de acordo com o que as pessoas farão com ele, como ele será utilizado (PEIRCE apud HJØRLAND, 1998) e isso não pode ser previsto. Assim, dois termos que no momento da elaboração do tesauro são realmente sinônimos, e a distinção entre eles naquele momento realmente não teria utilidade, com o passar do tempo podem deixar de ser, e a distinção vir a ser útil.

A forma exata como isso poderá ser feito ainda é uma questão a ser resolvida, uma possibilidade seria a formação de *clusters* dos resultados para cada um dos termos possíveis, como já ocorrem em algumas bases de dados que agrupam os resultados por autor, ou por ano

da publicação, por exemplo. No caso dos sinônimos, o critério de agrupamento seriam os diferentes termos que na busca inicial foram assim considerados.

Pode-se pensar que isso não seria necessário por que o usuário pode escolher o texto completo como campo de busca, opção presente em muitos sistemas, e assim fazer a pesquisa pelo termo que considera mais adequado, o que recupera apenas os documentos que contem tal termo no texto. Mas isso, em muitos casos, poderia resultar em baixa precisão, recuperando muitos documentos que apenas citam tal termo, tratando apenas brevemente o assunto.

Os profissionais da informação podem encontrar motivos para considerar que o que se propõe não seja realmente necessário, pois talvez já existam ferramentas que, devidamente combinadas poderiam oferecer o mesmo resultado. Descrever todas essas possibilidades ultrapassam o objetivo dessa pesquisa. Mas é importante lembrarmos que dominar o uso de tais ferramentas também pode ser bastante complexo, fazendo com que os usuários subutilizem os SRI.

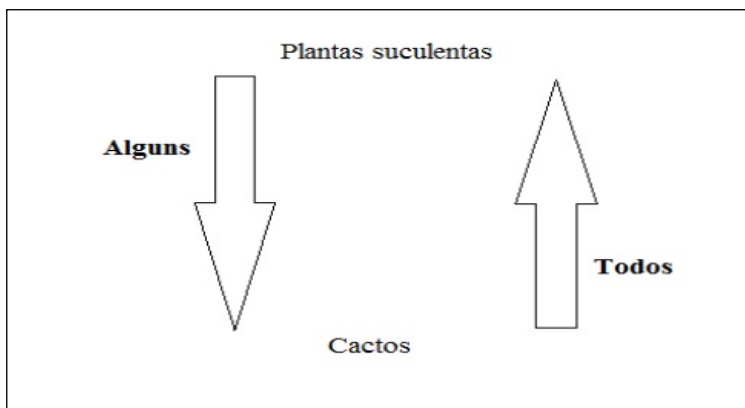
Essa sugestão de como pode ser feita a distinção dos termos que foram considerados equivalentes e formação de *clusters* é um exemplo de que há soluções que podem ser alternativas que levam em conta aspectos da abordagem pragmática. Estudos mais aprofundados são necessários, mas, nesta pesquisa foi possível identificar a importância dessa distinção.

4.3.2 Relações hierárquicas

As normas ISO25964-1(2011) e ANSI/NISO Z39.19 (2005) apontam a existência de três tipos de relações hierárquicas, gênero/espécie, todo/parte e de instância. Além desses tipos de relações hierárquicas, ambas as normas descrevem que um conceito pode ser logicamente designado como membro de mais de uma classe, o que se denomina relacionamento poli-hierárquico. A ISO25964-1(2011) também traz recomendações para a solução de problemas em tesauros multilíngues, que ocorrem quando os mesmos não possuem uma estrutura simétrica. A ANSI/NISO Z39.19 (2005) descreve também a possibilidade de se usar *node labels* para mostrar os princípios da divisão entre um conjunto de termos irmãos (que fazem parte de uma mesma classe).

Para o estabelecimento de relações hierárquicas do tipo gênero/espécie ambas as normas recomendam que se faça o teste “todos – alguns”, conforme esquema a seguir.

Figura 3 - Teste todos-alguns

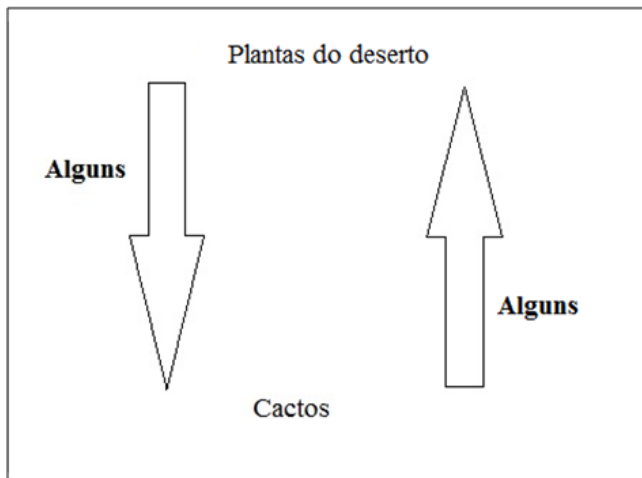


Fonte: ANSI/NISO Z39.19 (2005, p.47, tradução nossa)

A figura 3 mostra que alguns membros da classe plantas suculentas são conhecidos como cactos e que todos os cactos, por definição e, independentemente do contexto, são plantas suculentas.

A Figura 4 ilustra um caso do teste “todos – alguns” no qual alguns membros da classe plantas do deserto são conhecidos como os cactos e que alguns, mas não todos, os cactos são plantas do deserto. Assim, esses termos devem ser atribuídos a diferentes hierarquias no vocabulário controlado. (ANSI/NISO Z39.19, 2005).

Figura 4 - teste alguns-alguns



Fonte: ANSI/NISO Z39.19 (2005, p.48, tradução nossa)

A ISO25964-1(2011) apresenta o mesmo esquema para o teste mudando apenas os conceitos utilizados como exemplos, que no caso foram aves, papagaios e animais de estimação. Papagaios seriam hierarquicamente subordinados a aves, mas não a animais de estimação. Mas descreve também que:

Esse argumento não se aplica no contexto de um tesouro especialista dedicado a animais domésticos, em que os únicos papagaios no quadro de referência são animais de estimação. Nesse caso, “papagaios” pode ser subordinado a “animais de estimação” na mesma hierarquia. Tais aproximações devem ser aplicadas com cautela, especialmente em ambientes de rede em que os registros de um sistema podem ser misturados com os de outro. Quando é necessária a interoperabilidade, as relações estabelecidas devem ser universalmente aceitas. (ISO25964-1, 2011, p.59, tradução nossa)

Quanto a relação hierárquica todo-parte a norma ANSI/NISO Z39.19(2005, p.49, tradução nossa) descreve que:

Esta relação abrange situações em que um conceito é inerentemente incluído em outro, independentemente do contexto, para que os termos possam ser organizados em hierarquias lógicas, com o todo tratado como um termo mais amplo. Essa relação pode ser aplicada a alguns tipos de termo.

Apresenta três exemplos em que esse tipo de relação pode ocorrer: sistemas e órgãos do corpo, locais geográficos e estrutura hierárquica organizacional, empresarial, social ou política, e afirma que tais exemplos não são exaustivos.

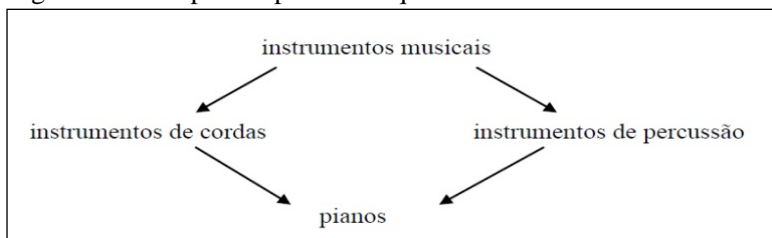
Porém na ISO25964-1(2011, p.60, tradução nossa) afirma-se que “relações todo-parte cobrem uma gama limitada de situações em que a parte de uma entidade ou sistema pertence exclusivamente a um todo particular”, essas situações ocorreriam apenas entre sistemas e órgãos do corpo, locais geográficos, disciplinas ou campos do discurso, estrutura social hierárquica.

Ambas as normas recomendam que quando um termo que representa a parte pertencer a mais de um todo, rodas, por exemplo, fazem parte de carros e bicicletas, não se deve implementar uma relação hierárquica todo-parte. Caso se considere apropriado, o tipo de relação que poderia ser estabelecida seria uma relação associativa.

A relação hierárquica de instância identifica a ligação entre uma categoria geral de coisas ou eventos, expresso por um substantivo comum, e uma instância individual dessa categoria, muitas vezes, um nome próprio. O exemplo apresentado para esse tipo de relação em ambas as normas é “Alpes” e “Himalaia” como termos subordinados à “regiões montanhosas”, que não são tipos nem partes de “regiões montanhosas” mas na verdade representam uma instância particular e são nomes próprios.

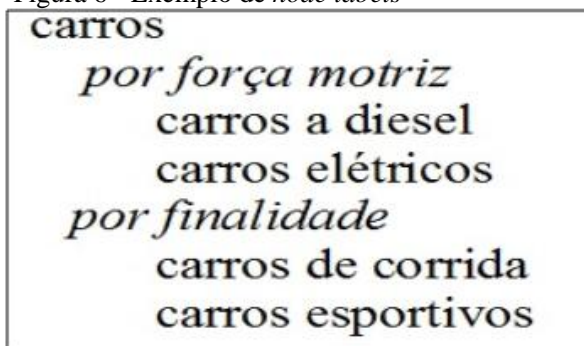
Quando um conceito pode ser logicamente designado como membro de mais de uma classe ou categoria um relacionamento poli-hierárquico também pode ser implementado. No exemplo a seguir, que é encontrado em ambas as normas, o termo “pianos” é subordinado a “instrumentos de cordas” e “instrumentos de percussão”, a poli-hierarquia nesse exemplo é baseada em relações hierárquicas gênero-espécie.

Figura 5 - Exemplo de poli-hierarquia



Fonte: ANSI/NISO Z39.19(2005, p.50) e ISO25964-1(2011, p.62, tradução nossa)

Além das recomendações referentes a essas possibilidades de relações hierárquicas, a norma ANSI/NISO Z39.19(2005) descreve que quando os termos são organizados em hierarquias *node labels* podem ser utilizados para mostrar os princípios da divisão entre um conjunto de termos irmãos termos (termos que compartilham o mesmo termo mais amplo). Alerta que, embora a sua função seja semelhante a de termos genéricos, *node labels* não são termos e não devem ser utilizados como termos de indexação.

Figura 6 - Exemplo de *node labels*

Fonte: ANSI/NISO Z39.19 (2005)

Node labels pode ser traduzido como indicador de faceta e é utilizado na apresentação de um tesauro quando se emprega a análise facetada. De acordo com a ANSI/NISO Z39.19 (2005, p.14) a análise facetada pode ser útil em tesauros que possuem milhares de termos, no qual uma estrutura puramente hierárquica pode se tornar confusa e a organização a partir da análise facetada pode facilitar seu manuseio.

A ISO25964-2(2011) aponta que a análise faceta é útil em gerar hierarquias conforme as regras para relações hierárquicas válidas apenas para conceitos que pertencem à mesma categoria geral. A escolha das facetas pode variar dependendo do domínio, mas de forma geral se usam as categorias fundamentais como objeto, matéria, agente, ação, espaço, tempo, etc. Caso se considere necessário pode-se subdividir essas facetas.

Diante dessas recomendações para o estabelecimento de relações hierárquicas que as normas apresentam, e que procuramos apresentar de forma sintetizada, juntamente com a análise da abordagem pragmática na OC, podemos dizer que o teste todos – alguns representa um exemplo claro de premissa universalista para o estabelecimento de relações semânticas. O que é criticado na abordagem pragmática, na qual se considera que as relações semânticas “devem ser entendidas principalmente quanto ao domínio específico, reveladas pelas (e construído pelas) disciplinas científicas” (HJØRLAND, 2003, p. 107). O teste “todos – alguns” é um teste de lógica referente à semântica formal, ou abordagem positivista em semântica, que conforme descrevemos, não precisa e nem deve ser ignorado, mas deve ser entendido assim como a semântica formal “como uma forma de destacar e materializar cada expressão contribuindo para as inferências de acordo com o contexto onde ocorrem” (PEREGRIN, 2004, p.12).

Nas recomendações das normas descreve-se que o teste todos – alguns não precisa ser aplicado em um tesauro de uma área específica, mas isso é descrito como uma espécie de exceção à regra, quando na verdade deveria ser o contrário. Até mesmo porque de forma geral praticamente todos os tesouros se destinam a domínios específicos. Assim, a resposta do teste todos – alguns poderia ter como base a análise da bibliografia do domínio ao qual o tesauro se destina, como critério para a consistência no estabelecimento de relações hierárquicas, mas um critério que é aplicado em contextos específicos, e não como um critério para relações hierárquicas universais.

Os outros dois tipos de relações hierárquicas, todo-parte e de instância, conforme descrevemos se destinam apenas a determinadas situações, que são devidamente explicadas e justificadas nas normas. Sobre as quais, com base na abordagem pragmática na OC, não temos maiores contribuições a fazer.

A análise facetada é uma técnica que vem sendo aplicada cada vez mais na OC, e no desenvolvimento de tesouros. Ela é mais flexível e adaptável do que demais regras aplicadas para o estabelecimento de

relações hierárquicas e pode ser usada como um item complementar para o desenvolvimento da estrutura hierárquica do tesauro.

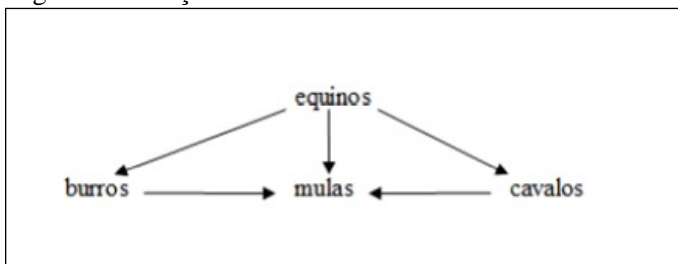
Em tesouros que contem milhares de termos, a estrutura puramente hierárquica pode se tornar confusa em alguns casos e dificultar o uso por parte dos usuários. Nesse sentido a análise facetada e o uso de indicadores de faceta vai ao encontro do que é defendido na abordagem pragmática na OC, que os tesouros e demais SOC devem servir como um mapa, um guia para o usuário encontrar a informação que necessita.

4.3.3 Relações associativas

Relações associativas cobrem ligações entre pares de conceitos que não são hierarquicamente relacionados, mas são semanticamente e conceitualmente associados de tal forma que a relação entre eles deve ser explicitada no tesauro, pois pode sugerir termos adicionais ou alternativos para uso na indexação ou RI. Como orientação geral para se estabelecer esse tipo de relação estaria o fato de sempre que um termo é usado, o outro deve estar implícito no quadro de referência compartilhado pelos usuários do tesauro. Além disso, um dos termos é frequentemente um componente necessário de qualquer explicação ou definição do outro. O termo células, por exemplo, é necessário para a definição de citologia, bem como o termo aves é um termo necessário para definição de ornitologia. (ANSI/NISO Z39.19, 2005; ISO25964-1, 2011).

A norma ANSI/NISO Z39.19(2005) define dois grupos principais de relações associativas, entre termos da mesma hierarquia e entre termos que pertencem a hierarquias diferentes. Entre os termos que pertencem a uma mesma hierarquia recomenda-se que a relação associativa não seja estabelecida entre termos irmãos mutuamente excludentes, por exemplo rosas e narcisos que pertencem à categoria flores. Mas pode ser implementada entre conceitos ligados por uma relação familiar ou derivacional (um conceito é derivado do outro), conforme o exemplo da figura 7:

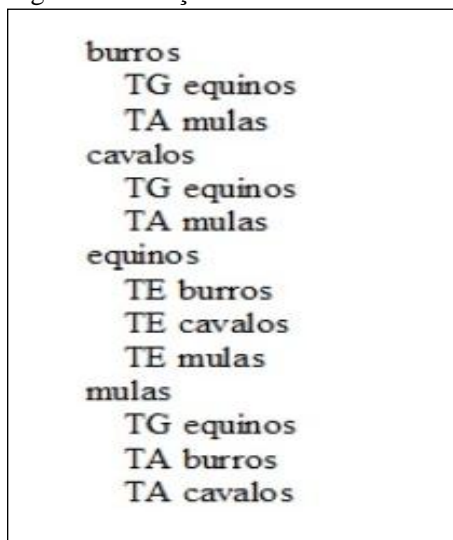
Figura 7 - Relação derivacional



Fonte: ANSI/NISO Z39.19(2005, p.52, tradução nossa).

Essa relação apareceria da seguinte forma no índice alfabético:

Figura 8 - Relação derivacional



Fonte: ANSI/NISO Z39.19 (2005, p. 52, tradução nossa).

Nos casos de termos pertencentes a uma mesma hierarquia, ou termos irmão, a norma ANSI/NISO Z39.19(2005) também recomenda que a relação associativa seja estabelecida entre aqueles que possuem significados sobrepostos. Por exemplo, barcos e navios, que podem ser definidos distintamente com precisão, sem formar um par de termos equivalentes, mas muitas vezes podem ser usados como sinônimos,

assim os usuários devem ser alertados ou lembrados, por meio da relação associativa entre eles.

A norma ISO25964-1 (2011) também recomenda o estabelecimento de relação associativa entre termos com significados sobrepostos, sendo este o critério de distinção dos dois grupos principais de relações associativas que essa norma apresenta, relação associativa entre termos com significados sobrepostos (que podem ou não pertencer a uma mesma categoria) e outros casos de ligações associativas.

Esse último grupo descrito na ISO25964-1 (2011), no item ‘outros casos de ligações associativas’, é similar ao grupo de relações associativas entre termos de diferentes hierarquias que a norma ANSI/NISO Z39.19(2005) descreve, no qual as relações associativas podem ser estabelecidas de acordo com variada gama de critérios, como por exemplo causa/efeito (morte/luto), objeto/propriedade (veneno/toxidade), ação/produto (tecelagem/tecidos) e processo/agente (controle de temperatura/termostato), dentre outros.

O que nos chama a atenção nessa recomendação é a definição de relações associativas, como aquelas que podem sugerir termos adicionais ou alternativos para uso na indexação ou RI. Essa descrição inicial vai ao encontro de uma importante discussão presente na abordagem pragmática, sobre o usuário ser informado da existência de pontos de vistas diferentes dos seus.

No entanto, a partir dessa definição, as normas descrevem como orientação geral para se estabelecer esse tipo de relação o fato de que sempre que um termo for usado, o outro deve estar implícito no quadro de referência compartilhado pelos usuários do tesauro. Com base na abordagem pragmática, podemos questionar como esse quadro de referência será identificado. Por exemplo, qual a finalidade de o tesauro trazer o termo ômega 3 se a coleção da unidade de informação não contém nenhum documento sobre tal assunto? Ao encontrar o termo no tesauro, os usuários podem criar a expectativa de encontrar material sobre o assunto, mas o resultado da busca será o silêncio.

Esse exemplo diz respeito à identificação de conceitos e de termos, mas ilustra de forma clara o que é defendido na abordagem pragmática, que os componentes dos SOC devem ser identificados a partir da bibliografia da área. Podemos complementar essa ilustração descrevendo, também como exemplo, que os usuários poderiam considerar que o ômega 3 é um ácido graxo presente na gordura de animais bovinos e suínos, mas na literatura indexada pelo tesauro, o ômega 3 é considerado um ácido graxo encontrado em peixes de águas profundas. Se uma relação associativa for implementada entre “ômega

3” e “animais bovinos” um usuário interessado em informações adicionais sobre o ômega 3, ao verificar tal relação pode decidir pesquisar também “animais bovinos” e não encontrará o que necessitava, pois na bibliografia o ômega 3 é considerado proveniente de peixes de águas profundas e a literatura sobre animais bovinos não conterá informações sobre ômega 3.

O que descrevemos é simplesmente a título de ilustração, sabemos que isso não acontece na realidade, os usuários possuem um certo conhecimento sobre a literatura da área. No entanto, essa literatura é muitas vezes bastante extensa e possui pontos de vista que não são compartilhados por todos os usuários. Desse modo, como a função das relações associativas, conforme é apontado inclusive nas normas para elaboração de tesouros, é sugerir termos adicionais para a indexação e RI e os termos representam conceitos que são definidos a partir de diferentes pontos de vista, a abordagem semântica pragmática se mostra especialmente útil para orientar o estabelecimento desse tipo de relação semântica. Essa abordagem enfatiza que as relações semânticas e demais componentes dos SOC devem ser identificados a partir da bibliografia da área. Pois um SOC deve servir como um guia para o usuário encontrar a informação que necessita. Essa informação pode ou não ser aquilo que o usuário desejava ou esperava encontrar quando iniciou sua busca, mas deve ser informação útil para que ele alcance seu objetivo.

4.3.4 Mapeamento das relações entre diferentes vocabulários

A ISO25964-2 (2013) apresenta recomendações para o mapeamento entre vocabulários dos diferentes tipos de relações semânticas, hierárquicas, associativas e de equivalência. O mapeamento de conceitos equivalentes é o mais importante e notoriamente o que recebe mais recomendações.

Na ISO25964-2 (2013, p. 25) recomenda-se que o mapeamento hierárquico seja estabelecido entre conceitos quando um é claramente subordinado ao outro. Assim como relações hierárquicas em um único tesouro, descritas na ISO25964-1 (2011), a relação hierárquica entre conceitos pode ser genérica, de instância, ou todo-parte, dependendo da situação.

A ISO25964-2 (2013, p. 25) recomenda que algum tipo de mapeamento associativo pode ser estabelecido entre conceitos quando os mesmos não forem mapeados como equivalentes nem hierárquicos, mas forem semanticamente associados. Considera-se que a linha que divide um mapeamento associativo de um mapeamento de equivalência

inexato é mal definida e subjetiva, mas pode ser feita pragmaticamente de acordo com o contexto em que os mapeamentos são usados, levando-se em conta, por exemplo, o interesse dos usuários, o volume de recursos relevantes a serem pesquisados, e a capacidade de recuperação e apresentação de distinções feitas no sistema.

O mapeamento de equivalência deve ser estabelecido quando conceitos correspondentes são encontrados em dois ou mais vocabulários diferentes. Ao contrário do estabelecimento de equivalência se dá entre dois termos em um tesauro monolíngue, quando um deles é designado como termo preferido e os outros como termos não preferidos, em um contexto entre vocabulários o mapeamento de equivalência é entre conceitos e não há diferença de *status* entre eles ou entre termos preferidos ou notações que os representam. Tradicionalmente os vocabulários incluem diferentes seleções de conceitos e as desenvolvem com diferentes níveis de especificidade. Portanto, várias situações de equivalência diferentes podem surgir, com diferentes soluções (ISO25964-2, 2013).

Em uma situação ideal, o vocabulário de destino contém um conceito idêntico no âmbito do conceito do vocabulário de origem. Um mapeamento de equivalência um-para-um pode ser imediatamente estabelecidos entre os conceitos. (ISO25964-2, 2013, p. 21, tradução nossa).

Um exemplo que a norma apresenta para esse caso, que é denominado equivalência simples, é telemóvel/telefone celular.

Alerta que o inverso nem sempre se aplica, ou seja, “termos preferidos idênticos em dois vocabulários diferentes não devem ser aceitos como equivalentes sem verificar o conceito subjacente”. (ISO25964-2, 2013, p. 21, tradução nossa). O exemplo usado para ilustrar esse caso é o termo operação, que em um tesauro militar e em um tesauro médico provavelmente terá significados diferentes. Quando os contextos são similares, as diferenças de escopo podem ser sutis, mas, ainda assim, quando termos idênticos são encontrados em diferentes vocabulários, um mapeamento de equivalência deve ser estabelecido apenas se o conceito subjacente for julgado equivalente.

Além da equivalência simples, pode ocorrer também a equivalência composta, quando:

Um conceito complexo que é incluído em um vocabulário com apenas um termo preferido pode ser representado em outro vocabulário pela combinação de dois ou mais termos/conceitos. (ISO25964-2, 2013, p. 22, tradução nossa).

Nesses casos o mapeamento da equivalência composta, conhecido como um-para-muitos, pode ser estabelecido entre os conceitos. Dois tipos diferentes de equivalência composta são normalmente encontrados, sendo conhecidos como “interseção” e “cumulativo”.

O mapeamento de equivalência entre dois conceitos encontrados em diferentes vocabulários ainda pode ser classificado de acordo com seu grau de aplicabilidade como exato, inexato e parcial. Um marcador pode ser usado para indicar o grau de aplicabilidade no contexto em que será utilizado (ISO25964-2, 2013).

O mapeamento de equivalência é considerado exato quando os conceitos podem ser usados de modo intercambiável entre todas as aplicações que podem ser previstas para o mapeamento (ISO25964-2, 2013).

O mapeamento de equivalência inexato ocorre quando conceitos correspondentes de dois ou mais vocabulários não são exatamente os mesmos. O problema é particularmente acentuado entre vocabulários que surgiram a partir de diferentes comunidades culturais, nos quais comumente ocorrem conceitos que podem ser equivalentes apenas em alguns contextos ou conceitos que podem ter escopos sobrepostos ou pequenas diferenças de conotação. Além disso, quando o mapeamento é feito entre tesouros e esquemas de classificação é comum encontrar classes com legendas que correspondem a um termo preferido mas não são exatamente equivalentes (ISO25964-2, 2013).

Na norma ISO25964-2(2013,p.27, tradução nossa), também descreve-se que:

Quando a equivalência inexata ocorre entre conceitos sobrepostos pode surgir a seguinte questão “Quão grande é a sobreposição?” esta questão é importante em aplicações nas quais o usuário precisa avaliar o provável benefício de seguir um mapeamento a fim de recuperar as informações mais relevantes sem recuperar muitos documentos irrelevantes. Estimar o grau de sobreposição também pode ajudar com a

classificação dos resultados da busca. Para tais aplicações, pode ser útil caracterizar a equivalência inexata como “sobreposição maior” ou “sobreposição menor.

A equivalência parcial, de acordo com a ISO25964-2 (2013, p.27, tradução nossa), ocorre quando:

A única diferença entre os conceitos semelhantes em diferentes vocabulários é que um é um pouco mais abrangente do que o outro. Ou seja, um dos conceitos tem um alcance apenas parcialmente equivalente para o outro.

Um dos exemplos para esse caso é monarcas/reis. É esclarecido, no entanto, que não há uma forma única de identificação de equivalência parcial. Em vez disso, uma escolha bastante subjetiva e difícil tem de ser feita entre as diversas possibilidades, como estabelecer um mapeamento hierárquico, um mapeamento de equivalência inexato, ou um mapeamento de equivalência composto cumulativo.

De acordo com a norma, essa e outras decisões, conforme já descrevemos, devem ser tomadas levando-se em conta, por exemplo, o interesse dos usuários, o volume de recursos relevantes a serem pesquisados, e a capacidade de recuperação e apresentação das distinções feitas no sistema ao qual o mapeamento se destina. Dependem de diversos aspectos, dentre esses a forma ou as formas como os mapeamentos serão utilizados na RI.

Segundo a norma ISO25964-2 (2013, p. 28, tradução nossa) “no contexto de recuperação da informação, existem duas etapas principais nas quais os mapeamentos entre vocabulários podem ser usados: a) como parte do processo de indexação ou b) no momento da pesquisa”. Dependendo do contexto de aplicação e da especificidade dos vocabulários alguns tipos de mapeamento são mais úteis do que outros. No anexo A apresentamos um quadro comparativo, encontrado na ISO25964-2 (2013), da aplicação de diferentes tipos de mapeamentos na indexação e no momento da pesquisa.

A partir das recomendações referentes ao mapeamento de relações semânticas entre diferentes vocabulários e da análise comparativa da ISO25964-2 (2013) como um todo com a abordagem pragmática na OC, que apresentamos na seção 4.2.3, podemos apontar que, ao se mesclar ou unir diferentes contextos em um só, as

características individuais consideradas, à princípio, importantes para o desenvolvimento de um SOC individual provavelmente sofrerão alterações. Analisar apenas os vocabulários de origem pode ser insuficiente e assim, a bibliografia como um todo que se visa abranger seria um aspecto contextual importante a ser analisado de acordo com a abordagem pragmática na OC.

5 CONSIDERAÇÕES FINAIS

As relações semânticas são elementos essenciais em tesauros e tem recebido menor atenção do que termos e conceitos nas pesquisas da área de CI. Configuram-se, portanto, em um objeto de pesquisa relevante. O estabelecimento de relações semânticas em tesauros pode ser feito de maneiras e de acordo com pressupostos teóricos diferentes, que nem sempre são apresentados de forma explícita, mas isso não significa que não existam. Assim, essa pesquisa teve como objetivo geral analisar como a abordagem semântica pragmática é ou pode ser adotada para o estabelecimento de relações semânticas em tesauros.

Visando alcançar esse objetivo geral o primeiro objetivo específico foi caracterizar a abordagem semântica pragmática no campo de OC. Os resultados referentes a esse objetivo são descritos na seção 4.1, na qual apresentamos as características da abordagem pragmática em nove enunciados temáticos que foram elaborados com base na análise dos 79 trabalhos que compõem o corpus da pesquisa.

Esses resultados foram utilizados para identificar a compatibilidade entre as características da abordagem pragmática e as recomendações das normas para elaboração de tesauros, o segundo objetivo específico da pesquisa, cujos resultados apresentamos na seção 4.2. Em relação a esse aspecto, notamos que apesar de encontrarmos nas normas alguns aspectos compatíveis com os princípios da abordagem pragmática, esses possuem menor importância, uma vez que não se configuram em elementos centrais para o desenvolvimento de um tesauro. Um exemplo dessa constatação é a recomendação quanto a inclusão de notas de escopo e notas históricas, que se mostraram compatíveis com a abordagem pragmática, mas nas normas para elaboração de tesauros constam com itens opcionais. Por outro lado, outros aspectos incompatíveis se configuram em elementos centrais para o desenvolvimento de um tesauro, por exemplo a definição de conceito como unidade de pensamento e uma maior preocupação com a representação (essencialismo) do que com a efetiva comunicação.

As normas registram o que é legítimo no momento em que são elaboradas, no entanto nossos valores e crenças se modificam a cada momento, o que muitas vezes nos dá a sensação de que a normalização “congela” um momento na nossa história. No intuito de atender as peculiaridades do momento e ser internacional a norma ISO acaba por tratar alguns aspectos de forma generalizada.

Uma vez que as normas apresentam as práticas já consolidadas no desenvolvimento de tesauros, podemos dizer que a abordagem

pragmática ainda é pouco utilizada para o desenvolvimento de tesouros e para o estabelecimento de relações semânticas. Assim, no último item dos resultados procuramos mostrar como a abordagem pragmática na OC pode ser adotada. Para isso sistematizamos um conjunto de orientações para o estabelecimento de relações semânticas em tesouros, com base nas recomendações das normas e nos enunciados da abordagem pragmática na OC, que apresentamos na seção 4.3.

Para as relações associativas destaca-se a importância da bibliografia para a identificação das relações que podem guiar o usuário a localizar informações úteis ao alcance de seus objetivos. A importância dada à bibliografia talvez não se aplique de forma tão efetiva no caso das relações de equivalência, nas quais a identificação da variedade de termos que podem representar um conceito é essencial e para isso os usuários, e não apenas a bibliografia, podem ser consultados. No caso das relações de equivalência percebemos que os aspectos contextuais envolvidos na construção do conhecimento, que são múltiplos e heterogêneos e acarretam implicações éticas para a OC, podem ser úteis para o desenvolvimento de tesouros e demais SOC que sejam mais orientados à comunicação do que à representação do conhecimento. Da mesma maneira, para a função de comunicação, destaca-se a importância da prática para a formação dos significados.

Um aspecto enfatizado na abordagem pragmática diz respeito à dificuldade em se identificar efetivamente os significados dos termos. Diferentes significados podem ser identificados por diferentes pessoas e comunidades ou, ainda, por uma mesma pessoa ou comunidade que, como o tempo, podem atribuir significados distintos aos termos. A relação de equivalência é importante para unir os diferentes termos em um único ponto de acesso de assunto de uma busca por determinado conceito. Possibilitar a distinção do que foi considerado igual ou equivalente para fins de revocação na RI é uma medida que minimiza as implicações éticas de se escolher uma forma de expressão como termo preferido e proporciona maior flexibilidade ao SOC e ao SRI. Flexibilidade que é uma característica da linguagem e tão importante para a comunicação, mas que os SOC, por vezes, tem como prioridade eliminar. O que, conforme descrevemos, é oposto ao que defende a abordagem pragmática na OC.

Nas relações hierárquicas, assim como nas de equivalência, o entendimento que os significados variam de acordo com o contexto é um aspecto importante, pois se as hierarquias são estabelecidas apenas a partir de premissas universalistas podem se tornar demasiadamente rígidas. O teste todos – alguns, um exemplo clássico da lógica formal e

do paradigma positivista em semântica, poderia ser utilizado como critério para a consistência no estabelecimento de relações hierárquicas, mesmo numa abordagem pragmática, contanto que se aplique ao que é aceito como verdade na bibliografia do domínio ao qual o tesouro se destina. Dessa maneira, é um critério aplicado em contextos específicos, e não apenas como um critério geral para identificar relações hierárquicas universais.

Com princípios da abordagem pragmática, os SOC podem contribuir para uma comunicação efetiva, ao possibilitar a compreensão dos diferentes pontos de vista presentes numa comunidade discursiva e não optar por um como sendo o melhor para representar uma realidade que é, por natureza, heterogênea. As relações semânticas poderiam desempenhar papel crucial nessa comunicação, ao permitir que o usuário possa partir do seu ponto de vista inicial e ir até aqueles defendidos pela frente de pesquisa sobre o tema.

Os SOC devem acompanhar a evolução dos saberes e fazeres, refletida também nos usos da linguagem que estariam presentes na literatura dos diferentes domínios. Nessa perspectiva, poderiam estar mais próximos dos estudos da Sociolinguística ou da Socioterminologia no estudo dos contextos de uso, nos quais, de acordo com a visão da abordagem pragmática, encontraríamos os conceitos e relacionamentos que comporiam a espinha dorsal dos SOC.

Na abordagem pragmática encontramos contribuições para o estabelecimento dos três tipos de relações semânticas dos tesouros, conforme descrevemos. Entre essas contribuições destacam-se aquelas que se referem às relações associativas, pois essas são consideradas as mais difíceis de definir e sobre as quais ainda não existe pesquisa suficiente para determinar suas bases teóricas, conforme apontam Marroni (2006), Austin (1993), Lancaster (1986) e Gomes (1984), dentre outros. Nessa abordagem, as relações associativas devem ser estabelecidas principalmente a partir da análise da literatura, de modo a levar em conta a existência de diferentes pontos de vista ali presentes, e servir como um guia para o usuário localizar informações possivelmente inesperadas e adicionais, mas ainda assim úteis para que alcance seu objetivo. Nessa perspectiva, podemos sugerir como pesquisas futuras a realização de estudos voltados ao desenvolvimento, identificação ou aplicação de ferramentas, metodologias ou mecanismos compatíveis com os princípios teóricos identificados para o estabelecimento das relações associativas, processo que integra e não pode ser separado do desenvolvimento de um SOC como um todo.

REFERÊNCIAS

ALMEIDA, M.; SOUZA, R.; FONSECA, F. Semantics in the Semantic Web: A Critical Evaluation. **Knowledge Organizatio**.v. 38, n. 3, 2011.

ANSI/NISO Z39.19:2005. **Guidelines for the construction, format and management of monolingual controlled vocabularies**. Bethesda: NISO, 2005.

ASHER, R. E; SIMPSON, J. M. Y. **The encyclopedia of language and linguistics**. Oxford: Pergamon Press, 1994. 10v.

AUSTIN, D. **Diretrizes para o estabelecimento e desenvolvimento de tesauros monolíngues**. Brasília: IBICT/SENAI. 1993. 83p.

BARDIN, L. **Análise de conteúdo**. Lisboa: Edições 70, 1979. 225p.

BARITÉ, M.; FERNÁNDEZ-MOLINA, J. C.; GUIMARÃES, J. A. C.;MORAES, J. B. E. Garantia literária: elementos para uma revisão crítica após um século. **TransInformação**, Campinas, v. 22, n. 2, maio/ago, 2010.

BATISTA, G. H. R. Redes de conceitos. **Perspect. cienc. inf.**, Belo Horizonte, v. 9, n. 1, p. 6-17, jan./jun. 2004.

BLAIR, D. C. Information Retrieval and the Philosophy of language. **ARIST**, v. 37, p. 2-50, 2003.

BOCCATO, V. R. C. Os sistemas de organização do conhecimento nas perspectivas atuais das normas internacionais de construção. **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 2, n. 1, p. 165-192, jan./jun. 2011.

BRÄSCHER, M. **Elaboração de tesauros**. Brasília. 2010.

CAFÉ, L.; BRASCHER, M. Organização do Conhecimento: Teorias Semânticas como base para estudo e representação de conceitos. **Inf. Inf.**, Londrina, v. 16 . n. 3. p. 25-51, jan./jun. 2011.

CAMPOS, M. L. A; GOMES, H. E; MOTTA, D. F. **Elaboração de tesauro documentário**: conceito. 2004. Disponível em: <<http://www.conexaorio.com/bit/tesauro/conceito.htm>>. Acesso em: 25 abr. 2013.

CAMPOS, L.; VENÂNCIO, L.S. O objeto de estudo da ciência da informação: a morte do indivíduo. **Inf. Inf.**, Londrina, v. 11, n. 1, jan./jun. 2006.

CAPURRO, R. Epistemologia e Ciência da informação. Tradução de Ana Maria Resende Cabral; Eduardo Wense Cdias; Isis Paim; Lígia Maria Moreira Dumont; Marta Pinheiro Aun e Mônica Erichsen Nassif Borges. In: V ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO. **Anais...**Belo Horizonte: s.n. 2003. Disponível em: <http://www.capurro.de/enancib_p.htm>. Acesso em: 26 ago. 2014.

CLARKE, S. G. D.; ZENG, M. L. From ISO 2788 to 25964: the evolution of thesaurus standards towards interoperability and data modeling. **Information Standards Quarterly**, v. 12, n. 1, 2012.

CURRÁS, E. **Tesauros**: linguagens terminológicas. Brasília: IBICT, 1995. 286p.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**. v. 7, n. 2, p. 101-07, 1978.

FIRTH, J. R. **Papers in linguistics; 1934-1951**. London: Oxford University Press, 1957. (reimp. 1969) 233p.

FROHMANN, B. Taking policy beyond Information Science: applying the actor network theory for connectedness: information, systems, people, organizations. In: ANNUAL CONFERENCE CANADIAN ASSOCIATION FOR INFORMATION SCIENCE, 23., 1995, Edmond, Alberta. **Anais...** Edmond, Alberta, 1995.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002. 171p.

GOMES, H. E. **Diretrizes para elaboração de tesauros monolíngues**. Brasília: IBICT, 1984. 70p.

GOMES, H. E. **Manual de elaboração de tesouros monolíngües**. Brasília: PNBIES, 1990.

GREEN, R.; BEAN, C.; MYAENG, S. H. **The semantics of relationships**: An interdisciplinary perspective. Dordrecht, NL: Kluwer. 2002.

HJØRLAND, B. Fundamentals of Knowledge Organization. **Knowledge Organizatio**. v. 30, n. 2. 2003.

HJØRLAND, B. Semantic and Knowledge organization. **ARIST**, v. 41, p. 367-405, 2007b.

ISKO. **International Society for Knowledge Organization**. Disponível em: <<http://www.isko.org/about.html>>. Acesso em: 28 out. 2012.

ISO 25964-1:2011. **Thesauri and interoperability with other vocabularies**. Part 1: Thesauri for information retrieval. Geneve: International Standard Organization, 2011.

ISO 25964-2:2013. **Thesauri and interoperability with other vocabularies**. Part 1: Interoperability with other vocabularies. Geneve: International Standard Organization, 2013.

KHOO, C; NA, J. Semantic relations in information science. **ARIST**, v.40, 157-228, 2006.

KUHN, T. S. **A estrutura das revoluções científicas**. 6. ed. São Paulo: Perspectiva, 2001. 257p. (Debates, 115).

LANCASTER, F. **Vocabulary control for information retrieval**. Washington, D.C.: Information Resources, 1986. 233p.

LYONS, J. **Semantics**. Cambridge: Cambridge University Press, 1977.

MARCONDES, D. Desfazendo mitos sobre a pragmática. **ALCEU**, v.1, n. 1, p. 38-46, 2000.

MARRONI, G. N. B. **Identificação e delimitação de relações associativas em tesauros: um estudo de caso na área do direito do trabalho**. 2006. 127 f. Dissertação (Mestrado em Ciência da Informação). Universidade de Brasília, Brasília, 2006.

MAZZOCCHI, F.; TIBERI, M.; SANTIS, B.; PLINI, P. Relational Semantics in Thesauri: Some Remarks at Theoretical and Practical Levels. **Knowledge Organizatio**. v. 34, n. 4. 2007.

PEREGRIN, J. Pragmatization of Semantics. In: TURNER, K. **The emantics/Pragmatics Interface from Different Points of View**. Elsevier, Amsterdam. 1999. p. 419-442.

PEREGRIN, J. Pragmatism und Semantik [Pragmatism and semantics]. In: Fuhrmann, A.; OLSSON, E. J. **Pragmatisch denken** [Thinking pragmatically]. p. 89-108. Frankfurt am Main, Germany: Ontos. 2004.

SOERGEL, D. The Arts and Architecture Thesaurus (AAT): A critical appraisal. **Visual Resources**, 10, 1995. p. 369-400. Disponível em: <http://www.dsoergel.com/cv/B47_short.pdf>. Acesso em: 25 abr. 2013.

SUNDIN, O.; JOHANNISSON, J. Pragmatism, neo-pragmatism and sociocultural theory: Communicative participation as a perspective in LIS. **Journal of Documentation**, v. 61, n. 1, 2005.

VEGA-ALMEIDA, R. L.; FERNÁNDEZ-MOLINA, J. C.; LINARES, R.. Coordenadas paradigmáticas, históricas y epistemológicas de la Ciencia de la Información; una sistematización. **Information Research**, v. 14, n. 2, p.1-20, June 2009. Disponível em: <<http://informationr.net/ir/14-2/paper399.html>>. Acesso em: 14 ago. 2014.

APÊNDICE A - BIBLIOGRAFIA CONSULTADA

ABRAHAMSON, J. A.; RUBIN, R. L. Discourse structure differences in lay and professional health communication. **Journal of Documentation**, v. 68, n. 6, 2012.

ALMEIDA, C. C. . A biblioteconomia e a ciência da informação na taxionomia das ciências de Charles Sanders Peirce. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 3, n.1, p. 1-19, 2005.

_____. Conceito como signo: elemento semiótico para análise e mediação da informação. **Scire** (Zaragoza), v. 18, p. 49-56, 2012a.

_____. Mediação como processo semiótico: em busca de bases conceituais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 5, p. 1-18, 2012b.

_____. **Peirce e a organização da informação**: contribuições teóricas da Semiótica e do Pragmatismo. Marília, 2009. 416f. Tese (Doutorado). Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília, 2009.

_____. Pragmatismo e Semiótica de Peirce na organização da informação. **Scire** (Zaragoza), v. 16, p. 15-20, 2010

_____. Sobre o pensamento de Peirce e a organização da informação e do conhecimento. **Liinc em Revista**, v. 7, p. 104-120, 2011.

_____. The Methodological Influence of Peirce s Pragmatism on Knowledge Organization. **Knowledge Organization**, v. 39, p. 204-215, 2012c.

ALMEIDA, M. B.; CEDON, B. V.; PINHEIRO, M. M. K.. Princípios metodológicos para a caracterização da dimensão pragmática de documentos no desenvolvimento de ontologias biomédicas. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.1, p. 105-117, jan./abr. 2012.

ALVARENGA, L. Bibliometria e arqueologia do saber de Michel Foucault: traços de identidade teórico-metodológica. **Ciência da Informação**, Brasília, v. 27, n. 3, p. 253-261, 1998.

ANDRADE, I. A. A. **As dimensões semântica e pragmática da Web e dos mecanismos de busca no ciberespaço**. 2012. 121f. Dissertação (Mestrado Profissional em Gestão da Informação) – Universidade Estadual de Londrina, Londrina, 2012.

ASSIS, J.; MOURA, M. A. A qualidade da informação na *web*: uma abordagem semiótica. **Inf. Inf.**, Londrina, v. 16 n. 3, p. 96 – 117, jan./jun. 2011.

BLAIR, D. C. Information Retrieval and the Philosophy of language. **ARIST**, v. 37, p. 2-50, 2003.

BROTHMAN, B. Perfect present, perfect gift: finding a place for archival consciousness in social theory. **Arch Sci**, v. 10, 2010.

CAMPOS, L. F. B.; VENÂNCIO, L. S. O objeto de estudo da ciência da informação: a morte do indivíduo. **Inf. Inf.**, Londrina, v. 11, n. 1, jan./jun. 2006.

CAVALERI, S. A. Are learning organizations pragmatic? **The Learning Organization**, v.15, n. 6, 2008.

DIEFENBACH, S.; HASSENZAHN, M. The dilemma of the hedonic – Appreciated, but hard to justify. **Interacting with Computers**, 23, 2011.

EKBIA, H. R.; EVANS, T. P. Regimes of Information: Land Use, Management, and Policy. **The Information Society**, v.25, p. 328–343, 2009.

FAINBURG, L.I. Information seeking and learning: a comparison of Kuhlthau's information seeking model and John Dewey's problem solving model. **New Library World**, v. 110, n. 9/10, 2009.

FROHMANN, B. Documentary ethics, ontology, and politic. **Arch Sci**, v.8, 2008a.

_____. Documentation Redux: Prolegomenon to (Another) Philosophy of Informatio. **LIBRARY TRENDS**, v.52, n. 3, 2004.

_____. Revisiting “what is a document?” **Journal of Documentation**, v. 65, n. 2, 2009.

_____. Rules of indexing: a critique of mentalism in information retrieval theory. **Journal of Documentation**, v.46, n.2, 1990.

_____. Subjectivity and Information Ethics. **JASIST**, v.59, n.2, 2008b.

_____. The power of images: a discourse analysis of the cognitive viewpoint. **Journal of Documentation**, v. 48, n. 4, 1992.

GONZALEZ DE GOMEZ, M. N. A Informação: dos Estoques às Redes. **Ciência da Informação**, Brasília, v. 24, n.1, p. 77-83, 1995.

_____. A reinvenção contemporânea da informação: entre o material e o imaterial. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 2, p. 01-21, 2009.

_____. A vinculação dos conhecimentos: entre a razão mediada e a razão leve. **Liinc em Revista**, v.1, n.1, março, 2005.

_____. Comentário ao artigo: “Hacia um nuevo paradigma en bibliotecologia”. **Transinformação**, Campinas, v. 8, n. 3, p. 44-56, set./dez. 1996.

_____. Da representação do conhecimento ao conhecimento da representação. **Ciência da Informação**, v. 22, n.3, p. 217-222, 1993.

_____. O caráter seletivo das ações de informação. **Informare** (Rio de Janeiro), Rio de Janeiro, v. 5, n.2, p. 7-31, 2000.

_____. Para uma reflexão epistemológica acerca da Ciência da Informação. **Perspectivas em Ciência da Informação**. V.6, n.1, p.5-18, jan./jun. 2001.

GRACIOSO, Luciana de Souza. **Filosofia da linguagem e Ciência da Informação: jogos de linguagem e ação comunicativa no contexto das ações de informação em tecnologias virtuais**. Rio de Janeiro, 2008. 176f. Tese (Doutorado em Ciência da Informação) – PPGCI – Universidade Federal Fluminense/IBICT.

_____. Justificação e a ação de informação no contexto da pragmática virtual. **Liinc em Revista**, v.6, n.2, setembro, 2010a. p. 286- 300.

_____. Language Philosophy in the context of knowledge organization in the interactive virtual platform. *Journal of Systemics, Cybernetics and Informatics*, v. 10, p. 64, 2012.

_____. Parâmetros teóricos para elaboração de instrumentos pragmáticos de representação e organização da informação na Web: considerações preliminares sobre uma possível proposta metodológica. **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 1, n.1, p. 138-158, 2010b.

GUASGUE, K. C. G. D. **O pensamento reflexivo na busca e no uso da informação na comunicação científica**. 2008. 240f. Tese (Doutorado) – Universidade de Brasília. 2008.

GUEDES, Roger de Miranda. **A abordagem dialógica na indexação social**. 2010. 186 f. **Dissertação** (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

HJØRLAND, B. Arguments for Philosophical Realism in Library and Information Science. **Library Trends**, v. 52,n. 3, inverno 2004.

_____. Citation analysis: A social and dynamic approach to knowledge Organization. **Information Processing and Management**.v. 49, 2013a.

_____. Concept theory. **JASIST**, v.60, n.8, p.1519-1536, 2009.

_____. Deliberate Bias in Knowledge Organization? **Advances in Knowledge Organization**, v. 11, p. 256-260, 2008a.

_____. Documents, Memory Institutions and Information Science. **Journal of Documentation**, v. 56, n. 1, January 2000.

_____. Empiricism, rationalism and positivism in library and information science. **Journal of Documentation**, v.61, n. 1, 2005.

_____. Epistemology and the Socio-Cognitive Perspective in Information Science. **JASIST**, v.53, n. 4, 2002a.

_____. Evaluation of an Information Source Illustrated by a Case Study: Effect of Screening for Breast Cancer. **JASIST**, v. 62, n. 10, 2011a.

_____. Evidence-Based Practice: An Analysis Based on the Philosophy of Science. **JASIST**, v. 62, n. 7, 2011b.

_____. Fundamentals of Knowledge Organization. **Knowl. Org.**, v.30, n.2, p.87-111, 2003.

_____. Information: Objective or Subjective/Situational? **JASIST**, v. 58, n. 10, 2007a.

_____. Information retrieval, text composition, and semantics. **Knowledge Organization**. V.25, n.1-2, 1998a.

_____. Is classification necessary after Google? **Journal of Documentation**, v.6, n.3, 2012.

_____. Semantic and Knowledge organization. **ARIST**, v.41, p. 367-405, 2007b.

_____. The Foundation of the Concept of Relevance. **JASIST**, v. 61, n.2, 2010.

_____. The Importance of Theories of Knowledge: Browsing as an Example. **JASIST**, v. 62, n. 3, 2011c.

_____. The Importance of Theories of Knowledge: Indexing and Information Retrieval as an Example. **JASIST**, v. 62, n. 1, 2011d.

_____. Theory and metatheory of information science: a new interpretation. **Journal of Documentation**, v. 54, n. 3, 1998 b.

_____. Towards a Theory of Aboutness, Subject, Topicality, Theme, Domain, Field, Content... and Relevance. **JASIST**, v.52, n. 9, 2001.

_____. User-based and Cognitive Approaches to Knowledge Organization: A Theoretical Analysis of the Research Literature. **Knowl. Org.**, v. 40, n.1, 2013b.

_____. What is Knowledge Organization (KO)? **Knowl. Org.**, v.35, n.2, p.86-101, 2008b.

HJØRLAND, B ; ALBRECHTSEN, H. Toward a new horizon in Information Science: domain-analysis. **JASIST**, v.46, n. 6, 1995.

HJØRLAND, B. ; CHRISTENSEN, F. S. Work Tasks and Socio-Cognitive Relevance :A Specific Example. **JASIST**, v.53, n.11, p.960-965, 2002.

HJØRLAND, B. PEDERSEN, K. N. A substantive theory of classification for information retrieval. **Journal of Documentation**, v 61, n. 5, 2005.

MOSTAFA, S. P.; NOVA CRUZ, D.V. Patchwork como princípio de produção e organização do conhecimento. **DataGramaZero** , v.12 n.1, 2011.

MOSTAFA, S. P. Charles Peirce, Gilles Deleuze e a Ciência da Informação. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.1, 2012.

MOSTAFA, S. P. Um banho de empirismo: de Hume/Deleuze ao empirismo radical de Bruno Latour. **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 1, n.1, 2010.

NG, K. B. The applicability of universal pragmatics in information retrieval interaction: a pilot study. **Information Processing and Management**, v. 38, 2002.

NOVELINO, Maria Salet Ferreira. A linguagem como meio de representação ou de comunicação da informação. **Perspect. Cienc. Inf.**, Belo Horizonte, v. 3, n. 2, p. 137 - 146, jul./dez.1998.

RENDÓN ROJAS, M. A. Hacia um nuevo paradigma em bibliotecologia. **Transinformação**, Campinas, v. 8, n. 3, p. 17-31, set./dez. 1996.

RANGLES, T. J.; BLADES, C. D.; FADLALLA, A. The Knowledge Spectrum. **International Journal of Knowledge Management**, v. 8, n. 2, 2012.

REDIGOLO, F. M. ; ALMEIDA, C. C. . Algumas contribuições da perspectiva filosófico-semiótica de Peirce para a análise de assunto. **Datagramazero** (Rio de Janeiro), v. 13, p. 1-17, 2012.

RODRIGUES, J.; CARICATTI, A. A pragmática no contexto da identificação de autoria de textos. **Ci. Inf.**, Brasília, v. 38, n. 1, p. 124-133, jan./abr. 2009.

SALDANHA, G. S. Entre a retórica e a filologia: do pragmatism ao humanismo na epistemologia da Ciência da Informação. **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 2, n. 1, jan./jun. p. 47-67 , 2011.

_____. 2008. Imago e vivência: uma reflexão filosófica sobre o essencialismo e o pragmatismo na ciência da informação. In IX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO ENANCIB, **Anais...** 2008, São Paulo: ANCIB, pp. 1-15.

_____. Tradições epistemológicas nos estudos de organização dos saberes: uma leitura histórico-epistêmica a partir da filosofia da linguagem. **Liinc em Revista**, v. 6, n.2. 2010. p. 300- 315.

_____. **Viagem aos becos e travessas da tradição pragmática da Ciência da Informação: uma leitura em diálogo com Wittgenstein.** 2008. 302f. Dissertação (Mestrado) – Universidade Federal de Minas Gerais. 2008.

SIQUEIRA, H. S. Sobre a natureza da tecnologia da informação. **Ci. Inf.**, Brasília, v. 37, n. 1, jan./abr. 2008

SOUSA, B. P.; ALMEIDA, C. C. Um olhar semiótico sobre o processo de indexação: a questão da representação e do referente. **Informação & Sociedade** (UFPB. Online), v. 22, p. 23-34, 2012.

SUNDIN, O.; JOHANNISSON, J. Pragmatism, neo-pragmatism and sociocultural theory: Communicative participation as a perspective in LIS. **Journal of Documentation**, v.61, n. 1, 2005.

THELLEFSEN, T. L; THELLEFSEN, M. M. Pragmatic semiotics and knowledge organization. **Knowledge Organization**, v. 31, n. 3, p. 177-187, 2004.

ANEXO A – Table 4 – Comparison of Mapping implementation in the conversion of index terms and search terms

Table 4 — Comparison of mapping implementation in the conversion of index terms and search terms

Type of mapping/ Example	Applied to index terms	Applied to search terms
Exact, simple equivalence mad cow disease =EQ bovine spongiform encephalopathy	The index term "mad cow disease" is converted to "bovine spongiform encephalopathy". No negative effects should occur, provided the controlled vocabulary is used effectively.	A search for "mad cow disease" is converted to "bovine spongiform encephalopathy". Both precision and recall are as they would have been if the collection had been indexed with the target vocabulary.
Intersecting compound equivalence women executives EQ women + executives	Both broader terms are added to the metadata. This tends to increase recall in all searches involving the terms "women" or "executives". However, users searching specifically for the concept of women executives will not easily separate the relevant items from others dealing both with executives of either sex, and with women in any role at all.	The search becomes "women AND executives". Recall is unaffected; precision is reduced. This search will retrieve documents that concern executives of either sex as well as women, for example a document about male executives harassing women staff.

Table 4 (continued)

Type of mapping/ Example	Applied to index terms	Applied to search terms
Cumulative compound equivalence inland waterways EQ rivers canals	Both narrower terms are added to the metadata. Unless Vocabulary A also has the narrower terms "rivers" and "canals", there will be a negative effect on precision (because a search for "canals" could retrieve items that were indexed with "inland waterways" despite dealing only with "rivers").	The search becomes "rivers OR canals". There is no negative effect on recall, but precision may be reduced if the user wants his search to apply at the generic level only.
Inexact simple equivalence potted plants ~EQ house plants	The index term "potted plants" is converted to "house plants". For a search on house plants, precision is reduced, as it will retrieve some items dealing with potted plants situated outdoors. Recall too may be reduced, if Vocabulary A does not allow for the concept "house plants", since items dealing with non-potted house plants will not have been indexed with either term.	The search term "potted plants" is converted to "house plants". Precision is reduced, as the search may retrieve some items dealing with house plants that are in containers other than pots.
Hierarchical mapping (narrower to broader) streets BM roads	The index term "streets" is converted to "roads". Any negative effect is likely to be small, provided the controlled vocabulary is used effectively.	The search term "streets" is converted to "roads". Recall may be increased slightly and precision reduced, as the search has been broadened.
Hierarchical mapping (broader to narrower) roads NM streets	The index term "roads" is converted to "streets". This will have a negative effect on the precision of searches for "streets" as they will now retrieve documents dealing with various types of roads (e.g. motorways and country lanes) that would not normally be considered streets.	The search term "roads" is converted to "streets". Precision will not suffer but recall will be reduced, since only one type of road will be retrieved.
Associative mapping e-learning RM distance education	In this case the index term "e-learning" is converted to "distance education". As a result, users searching for distance education can expect to retrieve some relevant documents, together with some that discuss other applications of e-learning.	In this case the search term "e-learning" is converted to "distance education". A negative influence on both recall and precision can be expected.